Canada

N° 81-595-MIF au catalogue — N° 005

Éducation, compétences et apprentissage
Documents de recherche

Établir le lien entre les évaluations provinciales des élèves et les évaluations nationales et internationales

par Fernando Cartwright, Diane Lalancette, Jerry Mussio, et Dehui Xing,

Division de la Culture, tourisme et centre de la statistique de l'éducation
2001 Immeuble principal, Ottawa, K1A 0T6
Téléphone : 1 800 307-3382 Télécopieur : 1 613 951-9040

Toutes les opinions émises par les auteurs de ce document ne reflètent pas nécessairement celles de Statistique Canada ou le Ministère de l'Éducation de la Colombie-Britannique

Statistique Canada Statistics Canada

Ministère de l'Éducation de la Colombie-Britannique
British Columbia Ministry of Education

Éducation, compétences et apprentissage
Documents de recherche

Établir le lien entre les évaluations provinciales des élèves et les évaluations nationales et internationales

Fernando Cartwright, Statistique Canada
Diane Lalancette, Ministère de l'Éducation de la Colombie-Britannique
Jerry Mussio, Statistique Canada
Dehui Xing, Ministère de l'Éducation de la Colombie-Britannique

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2003

Tous droits réservés. Il est interdit de reproduire ou de transmettre le contenu de la présente publication, sous quelque forme ou par quelque moyen que ce soit, enregistrement sur support magnétique, reproduction électronique, mécanique, photographique, ou autre, ou de l'emmagasiner dans un système de recouvrement, sans l'autorisation écrite préalable des Services de concession des droits de licence, Division du marketing, Statistique Canada, Ottawa, Ontario, Canada K1A 0T6.

Mai 2003

N° 81-595-MIF2003005 au catalogue

Périodicité : Irrégulier

ISSN 1704-8893

ISBN 0-662-89054-X

Ottawa

This publication is available in English upon request (Catalogue no. 81-595-MIE2003005).

Statistique Canada

Ministère de l'Éducation de la Colombie-Britannique

Les opinions exprimées par les auteurs de ce document ne reflètent pas nécessairement celles de Statistique Canada, ou le Ministère de l'Éducation de la Colombie-Britannique

105-03912/

Comment obtenir d'autres renseignements

Toute demande de renseignements au sujet du présent produit ou au sujet de statistiques ou de services connexes doit être adressée à : Services aux clients, Culture, tourisme et centre de la statistique de l'éducation, Statistique Canada, Ottawa, Ontario, K1A 0T6; téléphone : (613) 951-7608; sans frais : 1 800 307-3382; télécopieur : (613) 951-9040; ou courrier électronique : educationstats@statcan.ca.

Pour obtenir des renseignements sur l'ensemble des données de Statistique Canada qui sont disponibles, veuillez composer l'un des numéros sans frais suivants. Vous pouvez également communiquer avec nous par courriel ou visiter notre site Web.

Service national de renseignements	1 800 263-1136
Service national d'appareils de télécommunications pour les malentendants	1 800 363-7629
Renseignements par courriel	infostats@statcan.ca
Site Web	www.statcan.ca

Renseignements sur les commandes

On peut se procurer ce produit n° 81-595-MIF2003005 au catalogue sur internet gratuitement. Pour obtenir un numéro de ce produit, les utilisateurs sont priés de se rendre à http://www.statcan.ca/cgi-bin/downpub/studiesfree_f.cgi.

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois, et ce, dans la langue officielle de leur choix. À cet égard, notre organisme s'est doté de normes de service à la clientèle qui doivent être observées par les employés lorsqu'ils offrent des services à la clientèle. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1 800 263-1136.

Remerciements

Nous tenons à remercier les élèves, les enseignants et les directeurs et directrices d'école qui ont participé à l'étude PISA de l'OCDE de 2000 et à l'Évaluation des habiletés de base (ÉHB) de la Colombie-Britannique de 2000. Nous exprimons aussi notre gratitude au personnel de la Student Assessment and Program Evaluation Branch du ministère de l'Éducation de la Colombie-Britannique et du Centre de la statistique de l'éducation à Statistique Canada qui ont assuré à ce projet un soutien continu. Nous avons apprécié tout particulièrement l'aide de Valerie Collins et de Markus Baer du Ministère de la Colombie-Britannique qui ont participé à l'analyse des instruments. Enfin, nous tenons à remercier Danielle Baum de Statistique Canada qui a travaillé à la préparation du document aux fins de publication.

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population, les entreprises, les administrations canadiennes et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques précises et actuelles.

Table des matières

Sommaire 5

1. Introduction 7

2. Aperçu de l'ÉHB et du PISA 9

3. Rendement en lecture en Colombie-Britannique selon l'ÉHB et le PISA 13

4. Établissement de liens entre l'ÉHB et le PISA 17

5. Résultats 20

6. Conclusions 26

Annexe A : Tableaux 27

Annexe B : Comparaison de l'ÉHB et du PISA 31

Annexe C : Méthodologie technique 33

Références 44

Notes 45

N° 81-595-MIF2003005 au Catalogue

Sommaire

La présente étude de faisabilité visait à élaborer des procédures techniques permettant aux ministères de l'Éducation d'établir des liens entre les tests provinciaux et les tests nationaux et internationaux afin de pouvoir comparer les normes et présenter les résultats selon une échelle commune.

Les procédures techniques ont été élaborées et utilisées pour relier les tests de lecture administrés dans le cadre de l'Évaluation des habiletés de base (ÉHB) (Foundation Skills Assessment ou FSA) qui a eu lieu annuellement en Colombie-Britannique, et le Programme international pour le suivi des acquis des élèves (PISA). Aux fins de la présente étude de faisabilité, nous avons examiné les liens entre le test de lecture de l'ÉHB administré aux élèves de 10e année au printemps de 2000 et le test de lecture du PISA administré à peu près au même moment à un échantillon de jeunes âgés de 15 ans.

Selon le lien établi entre les deux évaluations, la norme d'excellence en lecture de l'ÉHB en Colombie-Britannique est plus élevée que la norme du PISA utilisée pour identifier les meilleurs lecteurs dans 32 pays. Si l'équivalent de cette norme internationale était utilisé dans l'ÉHB pour identifier les lecteurs excellents, un plus grand nombre d'élèves de 10e année de la C.-B. seraient classés comme excellents dans l'évaluation provinciale annuelle. Par exemple, selon l'ÉHB 2000, 9 % des élèves de 10e année de la C.-B. ont dépassé les attentes. Le nombre d'élèves classés comme dépassant les attentes en 2000 aurait pratiquement doublé si l'ÉHB utilisait la norme équivalente du PISA pour identifier les lecteurs excellents, .

Le lien établi permet également d'examiner le rendement des élèves âgés de 15 ans relevant d'autres instances scolaires par rapport à l'évaluation menée en C.-B. Par exemple, la Finlande et l'Alberta ont enregistré les scores PISA les plus élevés de tous les pays et toutes les provinces en 2000 et ils continuent d'afficher un bon rendement lorsque leurs scores internationaux sont transposés à l'échelle de lecture utilisée en C.-B. Plus particulièrement, 81 % des élèves âgés de 15 ans de la Finlande répondent aux attentes en lecture de la C.-B. ou les dépassent, suivis de 80 % des élèves de 15 ans de l'Alberta. La proportion des élèves qui répondent aux attentes ou les dépassent est à peu près la même pour le Québec et la C.-B., soit de 77 % et de 76 %, respectivement. Dans l'ensemble du Canada, 75 % des élèves âgés de 15 ans répondent aux attentes en lecture de la C.-B. ou les dépassent.

Dans les pays du G-7, le pourcentage des élèves de 15 ans qui répondent aux attentes en lecture de la C.-B. ou qui les dépassent s'échelonne de 58 % en Allemagne, 59 % en Italie et 64 % aux États-Unis à 75 % au Canada au Japon.

Les estimations de la proportion des élèves dans une autre instance qui répondent aux attentes en lecture de l'ÉHB ou les dépassent sont généralement exactes, sur l'ensemble des secteurs de compétence, à trois points de pourcentage près 19 fois sur 20.

Les provinces peuvent utiliser la méthodologie élaborée aux fins de la présente étude pour comparer les normes provinciales avec les normes nationales et internationales afin de présenter les résultats selon une échelle commune. La méthodologie peut être utilisée pour établir des liens entre deux évaluations lorsque les deux tests sont administrés à un même échantillon d'élèves ou lorsque des échantillons équivalents sont sélectionnés de façon aléatoire. Les procédures sont valides pour relier les scores des tests au niveau du groupe (groupe de plus de 30 élèves, pour la plupart des statistiques); elles ne sont pas appropriées pour relier et présenter les scores au niveau individuel.

N° 81-595-MIF2003005 au Catalogue

1. Introduction

La plupart des ministères de l'Éducation au Canada administrent des tests dans les écoles élémentaires et secondaires à l'échelle de la province. Ces évaluations permettent aux ministères de faire état du rendement des élèves par rapport aux normes établies par les provinces.

En outre, tous les ministères participent maintenant systématiquement aux évaluations pancanadiennes et internationales de l'apprentissage des élèves. Les administrations provinciales participent à ces évaluations externes pour comparer le rendement de leurs élèves avec celui des élèves d'autres provinces et pays. Pour les responsables de l'élaboration des politiques partout au Canada, ces évaluations sont devenues d'importants outils qui leur permettent de juger dans quelle mesure les élèves sont bien préparés à prendre leur place dans la société mondiale du savoir.

Il est difficile de comparer les résultats des évaluations provinciales, nationales et internationales parce que les résultats aux tests sont consignés sur des échelles différentes.

Alors que les évaluations provinciales et internationales sont maintenant administrées systématiquement dans les écoles canadiennes, les rapports produits par les gouvernements provinciaux semblent souvent en contradiction avec ceux produits par les organismes internationaux. Il n'est pas exceptionnel, par exemple, pour les médias de signaler des « taux d'échec » de 30 % au dernier test provincial administré puis, quelques mois plus tard, d'annoncer que le score enregistré par la province est l'un des « meilleurs au monde » dans une évaluation internationale. L'un et l'autre reportages peuvent être corrects, mais ils peuvent être source de confusion pour ceux qui supposent que les résultats sont présentés selon une même échelle ou que les deux évaluations ont mesuré le même phénomène.

Même lorsqu'on peut montrer que les évaluations portent sur le même phénomène (p. ex., les compétences en lecture), comparer les résultats aux tests provinciaux et ceux des évaluations internationales équivaut un peu à comparer les rapports quotidiens sur la température dans les capitales dans le monde entier sans savoir quelle échelle, Fahrenheit ou Celsius, est utilisée dans les divers pays.

On peut comparer les résultats de différentes évaluations si on établit un lien entre leurs échelles et une échelle commune.

La présente étude de faisabilité visait à élaborer des procédures techniques permettant aux ministères de l'Éducation d'établir un lien entre les tests provinciaux et les tests nationaux et internationaux afin que les résultats aux tests puissent être présentés selon une échelle commune. Lorsque les instruments sont reliés, une province peut déterminer si ses normes sont plus élevées ou au contraire plus faibles que celles établies par des organismes nationaux ou internationaux; elle peut aussi faire état du rendement d'autres instances par rapport aux normes provinciales.

Il devient de plus en plus important de pouvoir comparer les résultats d'évaluations menées à grande échelle puisque les éducateurs et le public cherchent à concilier les différences perçues entre les rapports provinciaux, nationaux et internationaux sur le rendement des élèves.

La présente étude porte sur le lien établi entre l'Évaluation des habiletés de base (ÉHB) de la C.-B. et le Programme international pour le suivi des acquis des élèves (PISA).

Il y a une autre raison de relier les évaluations. Les évaluations à grande échelle sont coûteuses à élaborer et à administrer et les écoles doivent y consacrer beaucoup de temps et d'énergie. L'établissement de liens entre les tests provinciaux et internationaux promet d'accroître la richesse et la rentabilité des programmes d'évaluation provinciaux en permettant d'intégrer des repères internationaux aux rapports courants des évaluations provinciales.

Aux fins de la présente étude de faisabilité, nous examinons le lien établi entre les tests de lecture administrés dans le cadre de l'Évaluation des habiletés de base (ÉHB) de la C.-B. et ceux du Programme international pour le suivi des acquis des élèves (PISA) de l'Organisation de coopération et de développement économiques (OCDE). La présente étude est fondée sur les données d'évaluation des compétences en lecture recueillies dans le cadre de l'ÉHB 2000 et du PISA 2000, qui ont tous deux été administrés d'avril à mai 2000. Pendant cette période, 2 800 élèves de 10e année ont participé à l'ÉHB ainsi qu'au PISA, ce qui fournit une excellente occasion de relier les deux échelles d'évaluation.

Au printemps 2000, 2 800 élèves de 10e année de la C.-B. ont participé à l'ÉHB ainsi qu'au PISA, ce qui fournit une excellente occasion de relier les deux échelles d'évaluation.

Les procédures techniques utilisées aux fins de cette étude ont été élaborées par Statistique Canada et sont fondées sur les travaux d'autres chercheurs.

Le présent rapport résume les résultats de l'étude de faisabilité. La partie 2 fournit des renseignements de base sur l'ÉHB et le PISA. La partie 3 expose les résultats aux tests de l'ÉHB et du PISA pour les élèves de la C.-B. La partie 4 contient un aperçu des procédures utilisées pour relier les deux évaluations. La partie 5 expose les résultats de la procédure d'établissement des liens et des conclusions sont présentées à la partie 6.

N° 81-595-MIF2003005 au Catalogue

2. Aperçu de l'ÉHB et du PISA

L'ÉHB

L'Évaluation des habiletés de base est une évaluation annuelle menée à l'échelle de la province qui donne un instantané des compétences en lecture, écriture et mathématiques des élèves de la C.-B. Les habiletés évaluées sont liées au programme scolaire provincial et l'accent est mis tout particulièrement sur les compétences requises pour fonctionner à l'école et dans la vie de tous les jours.

L'ÉHB comprend un ensemble de tests portant sur les compétences en lecture, en écriture et en mathématiques qui sont administrés annuellement à tous les élèves admissibles de 4e, 7e et 10e année en C.-B.

Les tests de l'ÉHB sont administrés chaque printemps aux élèves de 4e, 7e et 10e année dans les écoles publiques et les écoles indépendantes financées par la province. Les résultats sont publiés au début de l'automne de l'année scolaire suivante. La première évaluation complète a eu lieu en 2000 et sert de référence pour les comparaisons des compétences en lecture et en mathématiques. Pour les compétences en écriture, l'année repère est 2001.

Le système scolaire de la C.-B. utilise les résultats de l'ÉHB pour surveiller le rendement des élèves et élaborer des plans en vue de l'améliorer.

L'ÉHB vise principalement à aider les écoles, les districts scolaires et le Ministère à évaluer dans quelle mesure les élèves acquièrent les habiletés fondamentales et à élaborer des plans en vue d'améliorer le rendement des élèves. Elle vise en deuxième lieu à fournir aux parents des renseignements supplémentaires sur les progrès de leurs enfants.

L'ÉHB est conçue et administrée par la Student Assessment and Program Evaluation Branch du ministère de l'Éducation de la Colombie-Britannique. Les enseignants et autres éducateurs participent activement à l'élaboration des items des tests, à la correction des items à réponse ouverte et à l'établissement des normes.

Dans l'ÉHB, les scores de lecture sont produits par application de la théorie des réponses aux items. Les résultats sont ensuite consignés selon trois normes de rendement : dépasse les attentes, répond aux attentes et ne répond pas aux attentes (voir l'encadré 1). Les scores requis pour chacune de ces normes de rendement ont été établis par le Ministère en 2000 après consultation d'un groupe d'experts. Un échantillon d'items de tests communs est utilisé pour relier les évaluations d'une année à l'autre de manière à ce que les résultats de l'ÉHB chaque année puissent être présentés selon une échelle commune. Le Ministère prévoit réexaminer les normes de l'ÉHB périodiquement.

ENCADRÉ 1

Normes de rendement de l'ÉHB

Dépasse les attentes : Le niveau de rendement de l'élève dépasse celui auquel un enseignant considérerait que l'élève répond à toutes les attentes établies à l'égard de cette année scolaire pour ce test. Le rendement de l'élève à ce test serait considéré comme excellent pour cette année scolaire.

Répond aux attentes : Le niveau de rendement de l'élève répond aux attentes établies à l'égard de cette année scolaire à ce test.

Ne répond pas aux attentes : En l'absence d'autres renseignements, il s'agit du niveau auquel l'enseignant voudrait en savoir davantage sur les raisons du faible rendement de l'élève.

Les rapports publiés de l'ÉHB mettent l'accent sur la proportion des élèves qui répondent aux attentes provinciales ou les dépassent.

Les résultats sommaires, qui sont fournis aux écoles et au grand public, montrent la proportion d'élèves à chaque niveau de rendement. D'autres données statistiques, comme les scores moyens, sont également fournies pour aider à interpréter les données, mais l'attention est davantage dirigée vers la proportion des élèves qui répondent aux attentes provinciales ou qui les dépassent.

Les résultats de l'ÉHB sont combinés avec d'autres indicateurs, comme les résultats portés aux bulletins scolaires et les taux de persévérance scolaire, pour fournir des renseignements clés aux fins des plans de croissance élaborés par les conseils de planification scolaire et des contrats de responsabilités préparés par les districts scolaires. Les écoles fournissent aux parents les résultats individuels de l'ÉHB pour leurs enfants ainsi que l'information explicative pertinente.

Le PISA

Le PISA évalue les compétences en lecture, mathématiques et sciences; il est administré en 2000, 2003 et 2006 à des échantillons d'élèves âgés de 15 ans dans 32 pays.

Le Programme international pour le suivi des acquis des élèves (PISA) représente un effort collectif des pays membres de l'Organisation de coopération et de développement économiques (OCDE). Le PISA évalue le rendement des élèves âgés de 15 ans dans trois domaines : la compréhension de l'écrit (lecture), la culture mathématique et la culture scientifique. On utilise le mot « culture » pour montrer que l'accent est mis sur les connaissances et les compétences requises pour participer à la société.

Dans le cadre du PISA, trois cycles sont prévus – 2000, 2003 et 2006 – chacun portant sur un domaine de la littératie. Le cycle de l'an 2000 portait principalement sur la compréhension de l'écrit, les deux autres aspects étant traités comme des domaines secondaires. Par conséquent, l'évaluation comprenait moins d'items portant sur les mathématiques et les sciences et ces items ont été administrés à un souséchantillon de participants. Les cycles de l'an 2003 et 2006 porteront respectivement sur la culture mathématique et la culture scientifique.

Le PISA vise à éclairer l'élaboration des politiques dans les pays membres de l'OCDE.

Le PISA évalue la mesure dans laquelle les élèves qui approchent de la fin de leur scolarité obligatoire ont acquis les connaissances et les compétences essentielles pour participer pleinement à la société. Il fournit des données sur le rendement des élèves en lecture, mathématiques et sciences et révèle quels facteurs influent sur l'acquisition de ces compétences à la maison et à l'école. Les résultats sont censés contribuer à l'élaboration des politiques.

N° 81-595-MIF2003005 au Catalogue

Comme dans le cas de l'ÉHB, les scores en lecture du PISA sont produits au moyen de la théorie des réponses aux items (bien que les deux évaluations utilisent des variantes différentes du modèle de réponse à l'item). Les résultats en lecture du PISA sont présentés au moyen de deux statistiques principales : les scores moyens et la proportion des élèves qui atteignent chacun des cinq niveaux de compétence en lecture (voir l'encadré 2). Les scores définissant chaque niveau de compétence sont établis par des spécialistes du domaine des pays participants et des spécialistes en mesure représentant l'OCDE. Comme dans le cas du programme d'évaluation de la C.-B., l'évolution des scores du PISA est mesurée au fil du temps, l'année 2000 étant l'année de référence.

Trentedeux pays ont participé à l'enquête PISA 2000. Au Canada, environ 30 000 élèves de 15 ans dans plus de 1 000 écoles y ont participé. Un grand échantillon a été tiré au Canada afin d'obtenir des résultats fiables aux niveaux national et provincial; on ne présente pas de résultats individuels pour les écoles ou les élèves.

L'enquête PISA 2000 comprenait également des questionnaires destinés à recueillir des renseignements de base auprès des élèves et des directions d'école. Au Canada, on a également communiqué avec les parents des élèves par téléphone dans le cadre de l'Enquête auprès des jeunes en transition. On communiquera avec les élèves qui ont participé à cette enquête tous les deux ans. L'un des objectifs clés de la présente étude est de cerner les facteurs qui ont une incidence sur la poursuite d'études supérieures et le succès sur le marché du travail.

ENCADRÉ 2

Niveaux de compétence en lecture du PISA

Les élèves dont le rendement se situe au niveau 5 (plus de 625 points) sont capables de mener à bien des tâches de lecture complexes, notamment traiter des informations difficiles à trouver dans des textes non familiers, comprendre de manière approfondie des textes de ce type et en dégager les informations pertinents pour la tâche à accomplir, procéder à des évaluations critiques et élaborer des hypothèses, faire appel à des connaissances spécialisées et recourir à des notions qui peuvent être inattendues.

Les élèves dont le rendement se situe au **niveau 4** (553 à 625 points) sont capables d'effectuer des tâches de lecture difficiles, notamment localiser des informations enfouies dans un texte, dégager du sens de nuances de langage et évaluer un texte de manière critique.

Les élèves dont le rendement se situe au **niveau 3** (481 à 552 points) sont capables d'effectuer des tâches de lecture de complexité modérée, notamment localiser de multiples fragments d'information, établir des liens entre différentes parties de texte et mettre l'écrit en relation avec des connaissances familières de la vie courante.

Les élèves dont le rendement se situe au **niveau 2** (408 à 480 points) sont capables d'effectuer des tâches de lecture élémentaires, notamment localiser des informations directes, établir divers types d'inférences d'un degré limité, découvrir le sens d'un passage bien défini de texte et utiliser des connaissances différentes pour le comprendre.

Les élèves dont le rendement se situe au **niveau 1** (335 à 407 points) sont uniquement capables d'effectuer les tâches de lecture les moins complexes des épreuves PISA, notamment localiser un fragment unique d'informaiton, identifier le thème principal d'un texte ou établir une relation simple avec des connaissances de la vie courante.

Les élèves dont le rendement est **inférieur au niveau 1** (moins de 335 points) ne sont pas capables de mettre couramment en œuvre les connaissances et les compétences les plus rudimentaires PISA cherche à mesurer. Ces élèves éprouvent de sérieuses difficultés à utiliser la lecture comme un outil pour étendre et améliorer leurs connaissances et leurs compétences dans d'autres domaines.

Les résultats sommaires sont communiqués aux pays membres de l'OCDE et aux médias mondiaux. Les scores moyens obtenus par chaque pays sont présentés, de même que la proportion des élèves qui atteignent chaque niveau de compétence en lecture. Au Canada, les résultats de l'enquête PISA 2000 pour chaque province et pour des groupes démographiques clés ont été communiqués au moment de l'annonce des résultats internationaux.

Les pays qui participent à l'enquête PISA utilisent les résultats de différentes façons. En Allemagne, les résultats de PISA 2000 ont abouti à un examen complet du système éducatif. Au Canada, les ministres provinciaux se fondent généralement sur les résultats du PISA, les données d'autres tests pancanadiens et internationaux et les données des tests provinciaux pour informer le public des progrès dans le domaine de l'éducation.

N° 81-595-MIF2003005 au Catalogue

3. Rendement en lecture en C.-B. selon l'ÉHB et le PISA

Résultats de l'ÉHB

Selon l'EHB 2000, 79 % des élèves de 10e année de la C.-B. répondent aux attentes provinciales en lecture ou les dépassent et 21 % des élèves ne répondent pas à ces attentes.

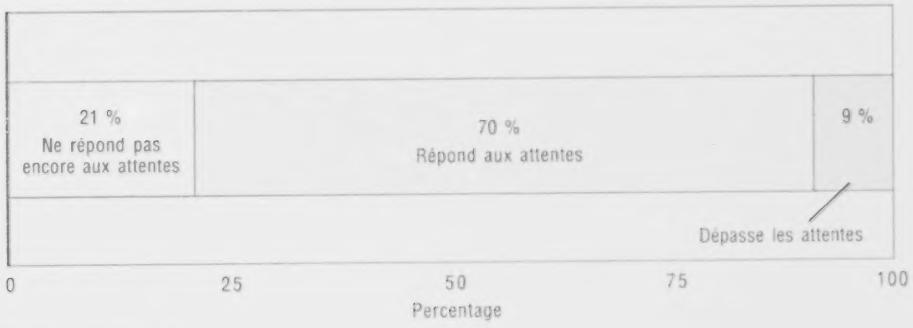
Au printemps 2000, les tests ÉHB ont été administrés aux élèves de 4e, 7e et 10e année. Tous les élèves devaient participer à l'évaluation; des directives permettant de dispenser certains élèves ayant des besoins spéciaux ont été fournies aux directions d'école chargées des décisions en matière d'exemption. Presque tous les élèves (94 %) de 4e et 7e année et 82 % des élèves de 10e année ont subi le test de lecture de l'ÉHB. Les élèves qui ont participé à l'ÉHB étaient ceux qui n'étaient pas exemptés de l'évaluation et ils sont donc considérés comme constituant une population de recensement.

La présente étude porte principalement sur le test de lecture de l'ÉHB administré aux élèves de 10e année au printemps de 2000. Selon les résultats de l'ÉHB, 79 % des élèves de 10e année ont répondu aux attentes provinciales ou les ont dépassées tandis que 21 % des élèves n'ont pas répondu aux attentes provinciales.

Figure 1

Pourcentage d'élèves de 10e année de la Colombie-Britannique répondant aux normes de lecture provinciales

Données de l'Évaluation des habiletés de base (ÉHB), 2000.



Source : Tableau 1.

Les résultats sommaires de l'ÉHB 2000 ont été communiqués aux 60 districts scolaires de la province et à toutes les écoles publiques ainsi qu'aux écoles indépendantes financées par la province. Le ministère de la C.-B. a publié non seulement les résultats obtenus par la population d'élèves, mais les résultats aux tests pour divers sousgroupes d'élèves en soulignant les variations dans le rendement. Le rendement en lecture des garçons et des élèves autochtones a été une source

d'inquiétude particulière en C.-B. Selon les résultats de l'ÉHB 2000, 74 % des élèves de sexe masculin de 10e année ont répondu aux attentes en lecture ou les dépassaient, comparativement à 85 % des élèves de sexe féminin de 10e année. 58 % des élèves autochtones ont répondu aux attentes en lecture ou les dépassaient comparativement à 79 % de la population dans son ensemble.

Entre 2000 et 2002, la proportion des élèves de 10e année qui ont répondu aux attentesenlecture de l'ÉHB ou qui les dépassaient a diminué, passant de 79 % à 71 %. La proportion des élèves de 7e année qui répondaient aux attentes en matière de lecture ou qui les dépassaient a également baissé durant cette période de trois ans, passant de 81 % à 76 %. Cette proportion a légèrement augmenté toutefois chez les élèves de 4e année, passant de 79 % à 80 %. Sur le plan des mathématiques et des sciences, la proportion des élèves de 4e, 7e et 10e année qui répondaient aux attentes ou les dépassaient a augmenté par rapport aux années repères (Tableau 1).

Résultats du PISA

Selon PISA 2000, les scores moyens en lecture des élèves de 15 ans de la C.-B. sont sensiblement les mêmes que ceux dans les instances qui affichent les scores les plus élevés, soit l'Alberta et la Finlande.

Scores moyens : Lorsqu'on compare les scores moyens obtenus aux tests dans les pays membres de l'OCDE, on constate que les jeunes de 15 ans au Canada obtiennent un rendement en lecture très élevé par rapport à leurs homologues dans les 32 pays qui ont participé à l'enquête PISA 2000. Seule la Finlande a affiché des scores moyens plus élevés que le Canada. La Nouvelle-Zélande, l'Australie, l'Irlande et le Japon se sont classés à peu près au même niveau que le Canada. (Tableau 2).

Les élèves de la Colombie-Britannique ont également obtenu des scores élevés comparativement à ceux enregistrés dans d'autres instances dans le monde. La figure 2 montre les scores moyens en lecture du PISA obtenus par les jeunes de 15 ans de la C.-B. comparativement à ceux de leurs homologues dans certaines provinces et dans les pays du G7[1]. La Finlande est également incluse dans la comparaison puisqu'elle a obtenu le score le plus levé des pays de l'OCDE dans la partie lecture de l'enquête PISA 2000.

Lorsqu'on tient compte des erreurs d'échantillonnage, les scores moyens enregistrés en C.-B. sont à peu près les mêmes que ceux affichés dans les instances qui ont enregistré les résultats les plus élevés, soit l'Alberta et la Finlande, et à peu près les mêmes que ceux enregistrés au Québec, en Ontario et au Japon. Les scores moyens obtenus par les élèves de la Colombie-Britannique sont plus élevés que ceux enregistrés au Royaume-Uni, en France, aux ÉtatsUnis, en Italie et en Allemagne ainsi que la moyenne pour l'OCDE.

Comme dans le cas de l'évaluation des compétences en lecture de la C.-B., dans l'enquête PISA, les filles de 15 ans ont obtenu des scores moyens en lecture plus élevés que leurs homologues de sexe masculin et ce, dans tous les pays participant au PISA. Une autre constatation clé découle d'une comparaison des scores moyens obtenus par les élèves de familles au statut socioéconomique (SSE) le plus élevé et de ceux obtenus par les élèves de familles au SSE le plus faible. Comparativement aux autres pays, le Canada et la Finlande affichent des résultats analogues, c'estàdire des scores élevés en lecture selon le PISA dans l'ensemble et de faibles écarts sur le plan du rendement selon le SSE.

N° 81-595-MIF2003005 au Catalogue

Figure 2

Note moyenne en lecture du PISA des jeunes de 15 ans en Colombie-Britannique et dans certaines instances

Données du Programme international pour le suivi des acquis des élèves (PISA) 2000

Les notes moyennes se situent à l'intérieur d'intervalles de confiance de 95 %.

Source : Tableau 2.

Niveaux de rendement : Le classement des scores moyens nous apprend peu de chose sur les capacités effectives des élèves. C'est pourquoi l'échelle du rendement en lecture dans l'enquête PISA 2000 a été divisée en cinq niveaux de rendement (encadré 2).

La figure 3 montre les profils de compétence en lecture pour certaines provinces et certains pays. L'Alberta et la Finlande compte une proportion élevée d'élèves au niveau supérieur de compétence en lecture soit 23 % et 19 %, respectivement, et une proportion faible au niveau inférieur, soit 8 % et 7 %, respectivement. En C.-B. également, une proportion élevée d'élèves se sont classés au niveau supérieur de compétence en lecture, soit 18 %, tandis que 9 % se sont classés aux niveaux inférieurs.

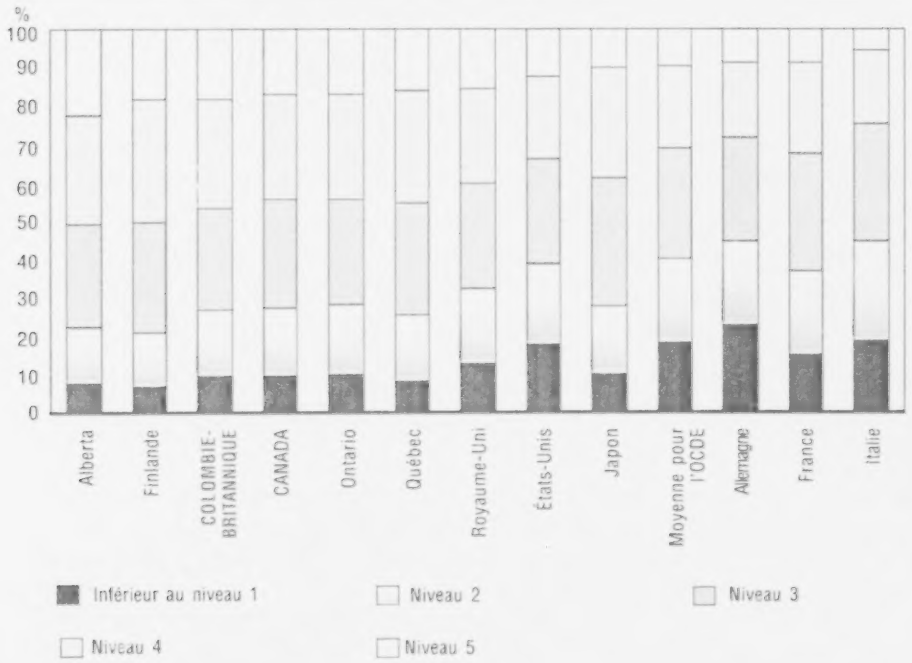
Selon les résultats de PISA 2000, 18 % des élèves de la C.-B. se classent au niveau le plus élevé de compétence en lecture sur le plan international et 9 %, au niveau le plus bas.

Dans la mesure où de fortes compétences en lecture permettent de prévoir une population et une main-d'œuvre plus instruites, les pays et les territoires de compétence qui comptent une forte proportion d'élèves qui se classent à un niveau de compréhension de l'écrit élevé jouiront d'un important avantage social et économique. Par contre, les jeunes dont les compétences en lecture sont faibles ou qui se classent à un niveau bas peuvent avoir de la difficulté à poursuivre d'autres possibilités de formation et être moins en mesure de contribuer et de participer aux sociétés de l'information et du savoir.

Figure 3

Pourcentage de jeunes de 15 ans qui se classent au niveau de compétence en lecture du PISA, C.-B. et certaines instances

Données du Programme international pour le suivi des acquis des élèves (PISA) 2000



Instances ordonnées selon le pourcentage d'élèves se classant au niveau 5

Source : Tableau 2.

Résumé

Selon les résultats de l'ÉHB 2000, 79 % des élèves de 10e année répondent aux attentes en lecture de la C.-B. ou les dépassent et 21 % ne satisfont pas à ces normes. Selon les résultats de PISA 2000, les jeunes de 15 ans de la C.-B., qui sont d'environ six mois plus jeunes en moyenne que les élèves de 10e année, enregistrent des scores élevés comparativement à leurs homologues dans d'autres pays. Selon les résultats du PISA, 18 % des jeunes de 15 ans de la C.-B. se classent au niveau supérieur de compétence en lecture sur le plan international et 9 %, aux niveaux inférieurs, résultats encore une fois très positifs au niveau international. Les normes de lecture en C.-B. sont-elles supérieures ou inférieures aux normes repères internationales établies par le PISA? Ces données provinciales et internationales peuvent-elles être présentées selon une échelle commune de manière à permettre de mieux comprendre les résultats?

N° 81-595-MIF2003005 au Catalogue

4. Établissement de liens entre l'ÉHB et le PISA

Comparaison des instruments

On a comparé l'ÉHB 2000 (lecture au niveau de la 10e année) et l'enquête PISA 2000 (lecture) pour déterminer si on mesure le même domaine de lecture.

Il n'est possible d'établir des liens entre les tests de lecture de l'ÉHB et du PISA que si ces tests visent à mesurer le même domaine de compétence ou de connaissance. Si les deux tests mesurent différents domaines de la lecture ou s'ils mesurent le même domaine de façon différente, alors les élèves afficheront probablement des tendances différentes en matière de rendement dans les tests; en pareil cas, les scores obtenus à un test ne permettront pas d'établir des estimations exactes des scores qui seront obtenus au deuxième test.

Dans le cas des élèves participant à l'ÉHB, on a analysé leur capacité de cerner et d'interpréter les concepts clés et les idées principales, leur habileté à repérer, interpréter et organiser des détails et leur capacité d'analyse critique. Dans le cas des élèves participant au PISA, on a évalué leur capacité de repérer de l'information particulière, d'interpréter le texte lu et d'y réfléchir et de l'évaluer en se fondant sur leurs connaissances acquises.

Pour l'ÉHB, on a utilisé un seul cahier de test comprenant quarante-et-un items de lecture administrés à tous les élèves; d'autres cahiers de test portant sur les composantes écriture et mathématiques ont été administrés durant la même semaine. Dans le cadre du PISA, 129 items ont été utilisés pour évaluer le rendement en lecture. Ces items ont été répartis en neuf cahiers et chaque élève faisant partie de l'échantillon a répondu aux questions figurant dans un cahier; huit des cahiers comprenaient également des items portant sur les mathématiques et les sciences.

Trois spécialistes de l'évaluation des capacités de lecture ayant une vaste expérience de la conception de l'ÉHB ont été priés d'examiner chacun des 41 items de tests de l'ÉHB, de classer chaque item selon le cadre d'évaluation du PISA et de cerner les items ne pouvant être classés. Ils ont ensuite été priés de procéder en sens inverse, c'estàdire d'examiner les 129 items de tests du PISA et de lier chaque item au cadre d'évaluation de l'ÉHB selon les mêmes critères. (Voir l'annexe A)

Les spécialistes ont conclu que les deux évaluations mesurent des habiletés de lecture similaires.

Les spécialistes ont signalé que les 41 items de test de l'ÉHB étaient classés dans le cadre d'évaluation du PISA. Dans le processus en sens inverse, les 129 items du PISA ont été classés dans le cadre de l'ÉHB. Les spécialistes ont signalé, toutefois, que 11 des items du PISA mettaient davantage l'accent sur l'application des compétences en lecture que ne le faisaient les critères de l'ÉHB. Dans l'ensemble, les spécialistes ont déclaré que les deux instruments mesuraient des habilités de lecture similaires. Ayant établi que les deux évaluations mesuraient à peu près le même phénomène, on a ensuite utilisé des procédures statistiques pour relier les échelles selon lesquelles les résultats étaient présentés.

Modèle statistique

Établir la relation mathématique entre les échelles de lecture de l'ÉHB et du PISA équivaut à déterminer la relation entre les échelles de température Fahrenheit et Celsius.

La prochaine étape du processus d'établissement d'un lien consiste à élaborer des modèles mathématiques reliant les deux évaluations. L'analogie avec les échelles de température aide à expliquer ce que suppose l'établissement d'un lien entre deux échelles de notation de tests. En analysant les mesures répétées de la température de l'air au moyen des deux thermomètres, Fahrenheit et Celsius, on peut établir la relation mathématique C = (5/9)(F-32). Cette formule montre qu'une température de 86 degrés sur l'échelle Fahrenheit équivaut à 30 degrés sur l'échelle Celsius. On peut analyser de la même façon les mesures du rendement des élèves en utilisant deux instruments – dans ce cas, le PISA et l'ÉHB – et utiliser les données pour établir la relation mathématique entre les deux échelles.

Cependant, il est beaucoup plus difficile de mesurer une activité cognitive comme la lecture que de mesurer la température et la formule qui établit un lien entre l'ÉHB et le PISA est plus complexe.

L'exemple de la température ne constitue qu'une analogie partielle puisqu'il est très simple de mesurer la température comparativement aux activités cognitives complexes comme la lecture ou les mathématiques. Bien que des erreurs puissent se produire si on utilise des thermomètres de mauvaise qualité, la température est mesurée directement et la formule reliant les échelles Celsius et Fahrenheit est exacte. La mesure d'une activité cognitive comme la lecture, cependant, est sujette à l'erreur due au fait que l'activité mesurée est souvent définie de différentes façons et mesurée indirectement, habituellement au moyen d'un test papier-crayon.

Pour relier les scores de l'ÉHB et ceux du PISA, il faut quantifier trois sources d'erreur : l'erreur de mesure, l'erreur de lien et l'erreur d'échantillonnage. L'erreur de mesure, qui décrit l'incertitude des scores originaux des élèves, est propre aux instruments de tests originaux. L'erreur de lien décrit l'incertitude qui accompagne la transposition d'un score sur une échelle originale en un score équivalent sur une autre échelle. Cette erreur, qui est exprimée sur la nouvelle échelle, est propre à la procédure mathématique utilisée pour estimer les scores équivalents, c'est-à-dire la fonction de lien. L'erreur d'échantillonnage décrit l'incertitude des statistiques utilisées pour estimer les caractéristiques (p. ex., les moyennes et les proportions) de la population dont un échantillon a été tiré. Une bonne partie des travaux techniques effectués dans le cadre de ce projet consistait à quantifier les deux premières erreurs et à les inclure dans la présentation des résultats de la fonction de lien. Les méthodes utilisées pour combiner les trois sources d'erreur dans la présentation de statistiques d'échantillon sont décrites à l'annexe C.

Trois stratégies peuvent être utilisées pour relier deux évaluations : inclure un échantillon des items de test communs dans les deux instruments; administrer les deux tests au même échantillon d'élèves; ou bien, sélectionner des échantillons équivalents de façon aléatoire. Les méthodes élaborées aux fins de la présente étude peuvent être appliquées aux deux dernières stratégies.

Les liens établis dans la présente étude étaient fondés sur les résultats aux tests de l'échantillon de 2 800 élèves de 10e année qui ont subi le test de l'ÉHB et le test du PISA durant une période d'un mois au printemps 2000. Même si les compétences en lecture des élèves pris individuellement peuvent évoluer durant cette période, les méthodes appliquées ici supposent que la répartition des compétences est similaire. En se fondant sur cette hypothèse, on a élaboré des modèles statistiques reliant les deux échelles d'évaluation. (La méthodologie technique utilisée pour établir les formules de lien est résumée à l'annexe C.)

N° 81-595-MIF2001005 au Catalogue

Bref, des différences significatives entre les éléments mesurés, la façon dont les tests ont été administrés et la fiabilité des deux instruments entraîneront inévitablement des erreurs d'interprétation. Dans la mesure où les deux tests portent sur les mêmes capacités, où ils ont été administrés dans les mêmes conditions à la même population et où les erreurs de mesure qu'ils comportent sont petites, l'erreur associée sera faible et les inférences tirées seront de forte validité.

5. Résultats

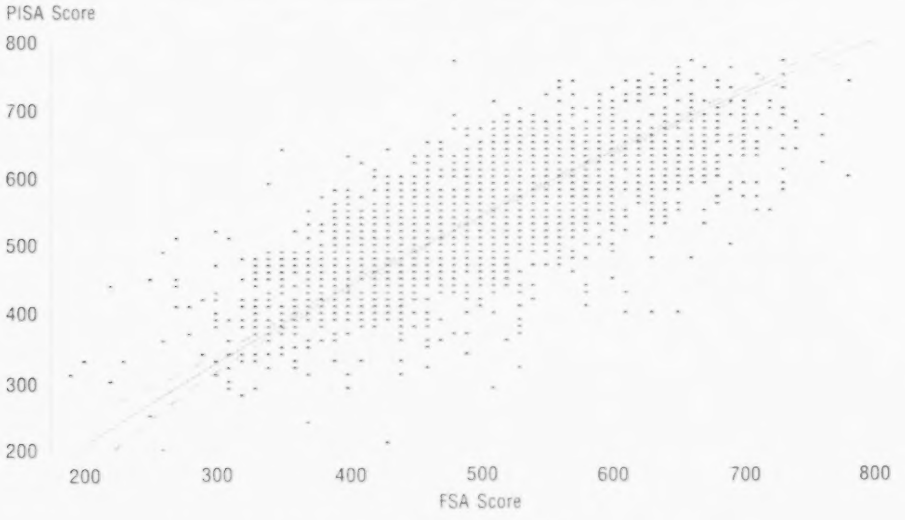
La relation mathématique entre l'ÉHB et le PISA est présentée à la figure 4.

On a élaboré des fonctions mathématiques reliant les deux évaluations, l'une reliant le PISA à l'ÉHB et l'autre, en sens inverse, reliant l'ÉHB au PISA. Les deux fonctions sont inverses l'une de l'autre. Par exemple, l'application d'une fonction à un score de l'ÉHB produira un score PISA équivalent, et l'application de l'autre fonction à ce score équivalent regénérera le score de l'ÉHB original.

L'équation mathématique exprimant le lien n'est pas aussi simple que la formule de conversion des degrés Celsius en degrés Fahrenheit, à cause de la complexité des procédures utilisées pour estimer la relation. La relation non linéaire entre les échelles ÉHB et les équivalents PISA est illustrée à la figure 4.

Figure 4

Rapport entre les notes de l'ÉHB et celles du PISA[1]



1. Les notes ont été arrondies à la dizaine près. Chaque point peut représenter plusieurs cas.
Les pointillés représentent l'intervalle de confiance de 68 % de la fonction de lien de l'ÉHB au PISA.

La fonction de lien est représentée sur un nuage de points représentants les paires de scores des élèves dans l'échantillon de lien, même si la méthode d'estimation n'a pas utilisé de méthodes de régression ou les paires de scores individuels (voir l'annexe C). Les lignes en pointillé représentent l'erreur-type de la fonction de lien de l'ÉHB au PISA. Le tableau 3, présenté à l'annexe A, montre un échantillon de

N° 81-595-MIF2003005 au Catalogue

scores du PISA et de scores de l'ÉHB liés ainsi que les erreurs de lien associées. Pour les scores faibles et élevés sur l'une et l'autre échelle, l'erreur est d'autant plus grande que l'incertitude concernant la valeur de l'équivalence des scores est importante. Dans le cas de la majorité des fourchettes des scores observés des deux échelles, on s'attend à ce que les scores équivalents soient exacts à 9 points près, 19 fois sur 20.

Les faibles différences entre les deux évaluations (relevées ci-dessus) laissent supposer que les scores reliés d'un test ne sont pas interchangeables avec les scores de l'autre test pour les élèves pris individuellement[2]. Toutefois, les hypothèses qui sous-tendent les méthodes d'établissement de liens exigent que les moyennes des scores reliés des groupes d'élèves soient interchangeables avec les moyennes des scores originaux. Par conséquent, il est valide de relier les scores de l'ÉHB et du PISA lorsque les résultats servent à tirer des inférences au sujet d'un groupe d'élèves (environ 30 élèves ou plus) mais non pour relier et présenter les scores au niveau individuel.

Comparaison des normes de lecture

Une fois les deux échelles reliées, les scores définissant les repères de lecture de l'ÉHB ont été convertis en scores du PISA et vice versa (tableaux 4, 5). À la figure 5, on compare les repères de l'ÉHB et du PISA au moyen de l'échelle internationale du PISA.

Le repère de l'ÉHB de l'excellence en lecture est établi bien au-dessus du seuil du niveau 5, le niveau le plus élevé de compétence en lecture du PISA.

Une comparaison des normes de l'ÉHB et de celles du PISA donne lieu à deux observations. En premier lieu, le seuil établi pour le niveau de rendement le plus élevé selon l'ÉHB, « dépasse les attentes », qui reconnaît l'excellence en lecture, est bien au-dessus du seuil établi pour le niveau 5 du PISA, qui est le plus haut niveau de compétence en lecture du PISA. Par conséquent, si l'ÉHB utilisait l'équivalent du niveau 5 du PISA comme norme pour identifier les meilleurs lecteurs, un plus grand nombre d'élèves de la C.-B. seraient classés chaque année comme dépassant les attentes provinciales. Par exemple, selon l'ÉHB 2000, 9 % des élèves de 10e année de la C.-B. dépassaient les attentes. Si l'ÉHB avait utilisé l'équivalent du niveau 5 du PISA comme norme pour identifier les lecteurs excellents, le nombre d'élèves classés comme dépassant les attentes en 2000 aurait doublé.

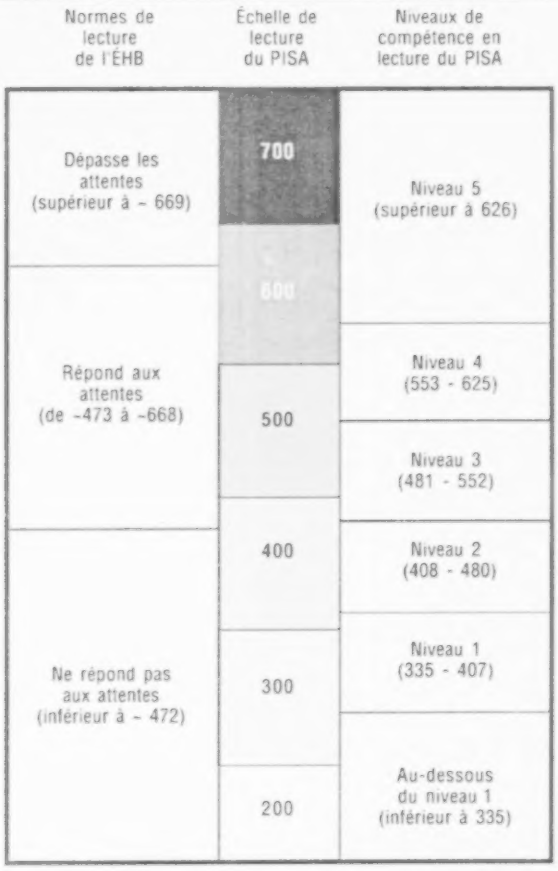
Si l'ÉHB utilisait l'équivalent du niveau 5 du PISA comme norme pour identifier les lecteurs excellents, un plus grand nombre d'élèves de 10e année de la C.-B. seraient classés dans l'évaluation provinciale annuelle comme dépassant les attentes.

En deuxième lieu, les scores de lecture classés dans la catégorie« répond aux attentes » dans l'ÉHB couvrent une large gamme de difficultés sur le plan de la lecture, soit les niveaux 3, 4 et 5 tels que définis par le PISA. On peut donc conclure que les élèves de la C.-B. classés dans la catégorie « répond aux attentes » sont capables d'accomplir des tâches de lecture de complexité modérée (niveau 3) mais peuvent aussi être capables d'accomplir des tâches de lecture plus complexes (niveaux 4, 5).

Les élèves de la C.-B. classés dans la catégorie « ne répond pas aux attentes » ont un rendement qui se situe au niveau 2 du PISA ou au-dessous. On peut conclure que ces élèves sont capables d'accomplir des tâches de lecture de base et de tirer des déductions élémentaires (niveau 2) ou qu'ils sont capables d'accomplir seulement les tâches de lecture les moins complexes (niveau 1), ou encore qu'ils ont beaucoup de difficultés à utiliser la lecture comme outil d'apprentissage efficace (rendement inférieur au niveau 1).

Figure 5

Comparaison des normes de l'ÉHB et du PISA selon l'échelle de lecture du PISA



Source : Tableaux 4, 5.

L'OCDE (2001) fournit des descriptions détaillées des habiletés associées à différents niveaux de rendement en lecture selon PISA.

ENCADRÉ 3

Score minimum ÉHB : Répond aux attentes	<— à peu près équivalents —>	Score minimum PISA : niveau 3
Niveau de rendement auquel l'élève répond aux attentes générales à l'égard de cette année scolaire à ce test.		Les élèves sont capables d'effectuer des tâches de lecture de complexité modérée, notamment localiser de multiples fragments d'information, établir des liens entre différentes parties de texte et mettre l'écrit en relation avec des connaissances familières de la vie courante.

N° 81-595-MIF2003005 au Catalogue

Comme on l'a mentionné plus haut, les normes de lecture de l'ÉHB et du PISA ont été établies après consultation d'un groupe de spécialistes. La C.-B. ou toute autre province peut choisir, pour différentes raisons, d'établir des normes supérieures ou inférieures à celles établies par d'autres provinces ou par des organismes nationaux ou internationaux comme l'OCDE.

Le niveau des normes établies dépendra des buts poursuivis par les tests. Dans la pratique, les normes de rendement dans la plupart des évaluations provinciales sont établies habituellement sans savoir si elles sont supérieures ou inférieures à celles établies pour des tests similaires administrés par d'autres instances. Si les évaluations provinciales sont reliées périodiquement aux tests externes, la province pourra comparer les normes et apporter des améliorations fondées sur une somme de connaissances plus complète.

Rendement des instances par rapport à la norme de lecture de la C.-B.

L'établissement de liens avec des tests externes permet à une province de déterminer les niveaux de rendement dans d'autres provinces et pays par rapport à ses normes provinciales. La figure 6 montre les résultats obtenus lorsque les scores du PISA sont reliés à l'ÉHB de la C.-B. Les résultats pour chaque instance sont exprimés par rapport à la norme de lecture de la C.-B.

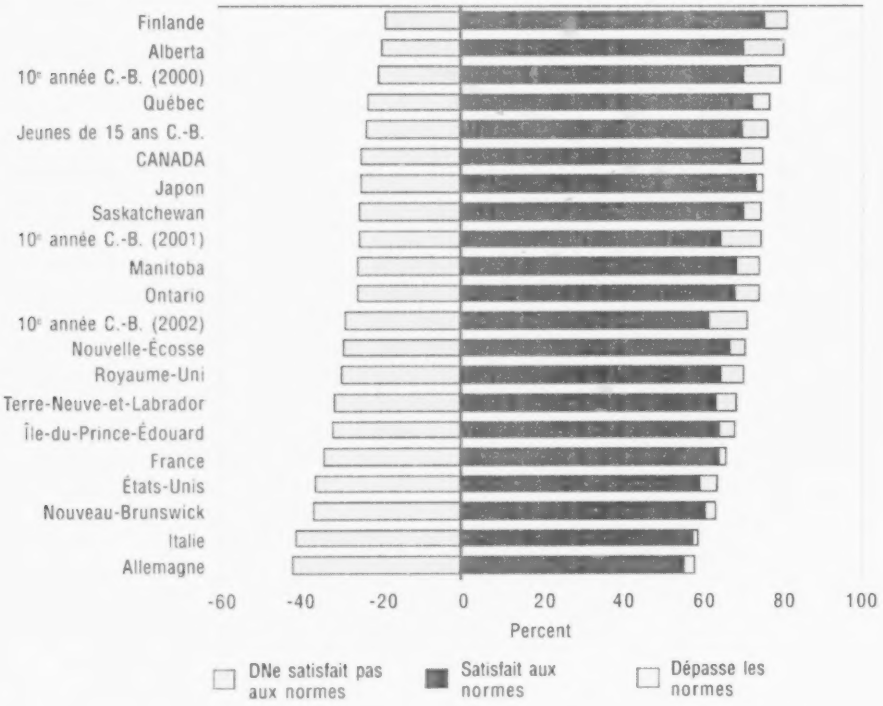
Relier les évaluations provinciales à des tests externes permet à une province d'intégrer des repères nationaux et internationaux dans ses rapports d'évaluation provinciale.

La Finlande et l'Alberta ont enregistré les scores du PISA les plus élevés de tous les pays et toutes les provinces en 2000 et ils continuent d'afficher un bon rendement lorsque leurs scores internationaux sont transposés à l'échelle de lecture utilisée en C.-B. Plus particulièrement, 81 % des élèves de 15 ans de la Finlande répondent aux attentes en lecture de la C.-B. ou les dépassent, suivis de 80 % des élèves de 15 ans de l'Alberta. La proportion des élèves qui répondent aux attentes ou les dépassent est à peu près la même pour le Québec et la C.-B., soit de 77 % et de 76 %, respectivement. Dans l'ensemble du Canada, 75 % des élèves âgés de 15 ans répondent aux attentes en lecture de la C.-B. ou les dépassent.

82 % des jeunes de 15 ans de la Finlande et 80 % de ceux de l'Alberta répondent aux attentes en lecture de la C.-B. ou les dépassent, comparativement à 77 % des élèves du Québec et à 76 % de ceux de la C.-B.

Figure 6

Pourcentage de jeunes de 15 ans dans diverses instances qui répondent aux normes de lecture de 10e année de la C.-B., 2000.



1. Tous les résultats sont ceux obtenus par les jeunes de 15 ans à l'exception des élèves de 10e année du secondaire de la C.-B. qui sont en moyenne de six mois plus âgés que les jeunes de 15 ans de la C.-B.

Les instances sont ordonnées selon le pourcentage d'élèves qui répondent aux attentes ou qui les dépassent.

Source : Tableau 6.

64 % des élèves de 15 ans aux États-Unis répondent aux attentes en lecture de la C.-B. ou les dépassent, comparativement à 75 % au Canada et au Japon.

Dans les pays du G-7, le pourcentage des jeunes de 15 ans qui répondent aux attentes en lecture de la C.-B. ou qui les dépassent s'échelonne de 58% en Allemagne, 59 % en Italie et 64 % aux États-Unis à 75 % au Canada au Japon.

Le tableau 6 comprend les erreurs-types associées aux statistiques qui sont produites lorsque les scores du PISA sont convertis en scores équivalents de l'ÉHB. Ces erreurs-types comprennent l'erreur de mesure, l'erreur de lien et l'erreur d'échantillonnage du PISA. Les estimations de la proportion des élèves d'une autre instance qui répondent aux attentes en lecture de l'ÉHB ou qui les dépassent sont généralement exactes, pour les différentes provinces et les différents pays, à 3 points de pourcentage près, 19 fois sur 20.

Il faut néanmoins interpréter ces résultats avec prudence. On a estimé ces liens au moyen d'un échantillon tiré de la population des élèves de 15 ans de 10e année de la C.-B.; au fut et à mesure que des différences systématiques entre cette population de lien et les régions à l'échelle desquelles ces résultats sont généralisés augmentent, la validité des interprétations diminue. Les pays dans lesquels le PISA

N° 81-595-MIF2003005 au Catalogue

a été administré diffèrent en ce qui concerne de nombreux indicateurs sociaux, économiques et culturels. On peut tirer des inférences solides pour des populations similaires à celle pour laquelle le lien a été estimé, comme les étudiants dans d'autres provinces canadiennes ou dans des pays qui se sont développés de façon analogue. Toutefois, les résultats de l'établissement des liens pour des populations qui sont très différentes de l'échantillon de lien doivent être interprétés comme étant des approximations grossières plutôt que des estimations précises.

Dans le cas des élèves de la C.-B. inscrits en 10e année en 2000, les résultats de l'ÉHB montrent que 79 % d'entre eux répondent aux attentes en lecture de la C.-B. ou les dépassent. Cette proportion baisse à 75 % en 2001 et à 71 % en 2002. Les élèves de 10e année de la C.-B. sont en moyenne de six mois plus âgés que les jeunes de 15 ans de la C.-B.

Comme on l'a mentionné ci-dessus, la norme de l'ÉHB reconnaissant l'excellence en lecture est plus élevée que la norme internationale du PISA utilisée pour identifier les meilleurs lecteurs. Par conséquent, un plus petit nombre d'élèves se classent au niveau supérieur de compétence en lecture (dépasse les attentes) dans l'ÉHB. Par exemple, 9 % des jeunes de 15 ans de la C.-B. se classent au niveau supérieur de compétence en lecture de l'ÉHB, tandis que 18 % se classent au niveau supérieur de lecture du PISA. Pour ce qui est des élèves de l'Alberta, 10 % se classent au niveau supérieur de l'ÉHB de la C.-B. et 23 %, au niveau supérieur de lecture du PISA.

Il convient de signaler qu'on peut aussi appliquer la formule d'établissement de liens de manière à représenter les scores de l'ÉHB pour les écoles et les districts scolaires de la C.-B. en fonction des niveaux de compétence en lecture du PISA.

6. Conclusions

Les évaluations provinciales, pancanadiennes et internationales sont maintenant couramment administrées dans les écoles canadiennes. Pour les responsables de l'élaboration des politiques dans tout le Canada, ces évaluations sont devenues d'importants outils qui leur permettent de juger dans quelle mesure les élèves sont prêts à prendre leur place dans une société mondiale du savoir.

Il a été difficile, cependant, de comparer les résultats des divers programmes d'évaluation parce que les échelles utilisées sont différentes. Il n'est pas exceptionnel pour les médias de signaler des taux d'échec élevés à un test provincial puis, quelques mois plus tard, de présenter les résultats positifs et, sembletil, contradictoires d'une évaluation internationale. La présente étude de faisabilité visait à élaborer des procédures techniques permettant aux ministères de l'Éducation d'établir les liens entre les tests provinciaux et les tests pancanadiens et internationaux afin de pouvoir comparer les normes de différentes évaluations et de présenter les résultats selon une échelle commune.

La présentation des résultats aux tests selon une échelle commune permettra au public de comprendre plus facilement le rendement des élèves de la province par rapport aux repères provinciaux et internationaux. Les évaluations sont coûteuses à concevoir et à administrer, de sorte qu'en reliant ces instruments et en augmentant la richesse des rapports d'évaluation on améliorera également la rentabilité des programmes d'évaluation provinciaux.

Les méthodes ont été élaborées et utilisées pour relier les tests de lecture administrés par l'Évaluation des habiletés de base (ÉHB) de la C.-B. et ceux du Programme international pour le suivi des acquis des élèves (PISA). Les normes de lecture de l'ÉHB ont été comparées aux normes du PISA. Les résultats du PISA pour les autres provinces et pays ont été transposés dans l'échelle utilisée en C.-B. et présentés par rapport aux normes de lecture de la C.-B.

On peut utiliser la méthodologie élaborée ici pour établir des liens entre deux évaluations lorsque les deux tests sont administrés à un même échantillon d'élèves ou lorsque des échantillons équivalents d'élèves sont tirés de façon aléatoire. On peut utiliser d'autres méthodes d'établissement de liens lorsque les évaluations sont conçues de manière à porter sur certains items communs ou lorsqu'elles sont administrées simultanément. On peut utiliser les programmes informatiques élaborés pour les besoins de la présente étude pour relier les échelles aux fins d'évaluations au moyen de diverses méthodes d'échantillonnage, y compris la théorie des réponses aux items et la théorie classique des tests.

Les liens de ce type sont valides pour relier des échelles de test lorsque les résultats sont utilisés pour tirer des inférences sur des groupes d'élèves (environ 30 élèves ou plus); ils ne sont pas appropriés pour relier et présenter les scores au niveau individuel.

N° 81-595-MIF2003003 au Catalogue

Annexe A : Tableaux

Tableau 1

Rendement des élèves de la C.-B. sur les tests de lecture, d'écriture et de mathématiques de l'ÉHB[1,2,3]
Évaluation des habiletés de base, 2000-2002

		Ne répond pas encore aux attentes	Répond aux attentes	Dépasse des attentes	Répond aux attentes ou les dépasse	Erreur-type
Lecture						
4e année	2000	20,5	71,6	7,9	79,5	(0,10)
	2001	22,1	72,8	5,1	77,9	(0,09)
	2002	20,1	73,5	6,4	79,9	(0,10)
7e année	2000	19,3	72,9	7,9	80,7	(0,10)
	2001	24,2	66,4	9,4	75,8	(0,10)
	2002	23,6	64,5	11,9	76,4	(0,10)
10e année	2000	20,6	70,5	8,9	79,4	(0,09)
	2001	25,3	64,7	10,0	74,7	(0,09)
	2002	28,8	61,7	9,6	71,2	(0,12)
Écriture						
4e année	2001	9,3	89,9	0,9	90,7	(0,13)
	2002	6,2	93,6	0,2	93,8	(0,11)
7e année	2001	19,0	77,9	3,1	81,0	(0,14)
	2002	15,6	82,8	1,6	84,4	(0,13)
10e année	2001	13,8	81,9	4,3	86,2	(0,13)
	2002	13,2	79,5	7,3	86,8	(0,14)
Calcul						
4e année	2000	20,5	71,1	8,4	79,5	(0,09)
	2001	16,1	73,4	10,5	83,9	(0,09)
	2002	14,6	71,9	13,6	85,4	(0,10)
7e année	2000	20,3	70,2	9,5	79,7	(0,09)
	2001	18,8	73,8	7,4	81,2	(0,10)
	2002	17,8	72,8	9,4	82,2	(0,11)
10e année	2000	25,3	65,9	8,7	74,7	(0,10)
	2001	23,3	67,1	9,5	76,7	(0,09)
	2002	23,9	64,9	11,2	76,1	(0,13)

1. Les erreurs-types présentées ici reflètent les erreurs de mesure des instruments de l'ÉHB.
2. Aucune erreur d'échantillonnage n'est calculée puisqu'on suppose que la population des étudiants a participé à l'ÉHB.
3. Pour l'évaluation des compétences en lecture et en mathématiques, l'ÉHB utilise une échelle de rendement dont la moyenne est égale à 500 et l'écart-type est égal à 100.

Tableau 2

Compréhension de l'écrit (lecture) du PISA : Moyennes et niveaux de compétence
Scores moyens et pourcentage d'élèves à chaque niveau de compétence sur l'échelle combinée de compréhension de l'écrit (lecture)
(Les erreurs-types sont indiquées entre parenthèses)

Instances[1,2]	Moyenne		Inférieur au niveau 1		Niveau 1		Niveau 2		Niveau 3		Niveau 4		Niveau 5	
Alberta	550	(3,3)	1,8	(0,5)	6,1	(0,7)	14,7	(0,8)	26,7	(1,2)	28,2	(1,0)	22,5	(1,4)
Finlande	546	(2,6)	1,7	(0,5)	5,2	(0,4)	14,3	(0,7)	28,7	(0,8)	31,6	(0,9)	18,5	(0,9)
Colombie-Britannique	538	(2,9)	2,4	(0,5)	7,0	(0,7)	17,5	(0,9)	26,3	(1,1)	28,7	(1,0)	18,1	(1,1)
Québec	536	(3,0)	2,0	(0,4)	6,4	(0,6)	17,2	(0,9)	29,4	(1,1)	29,2	(1,1)	15,9	(1,0)
CANADA	534	(1,6)	2,4	(0,3)	7,2	(0,3)	18,0	(0,4)	28,0	(0,5)	27,7	(0,6)	16,8	(0,5)
Ontario	533	(3,3)	2,6	(0,6)	7,4	(0,8)	18,2	(0,9)	27,5	(0,9)	27,6	(1,1)	16,7	(1,0)
Manitoba	529	(3,5)	2,0	(0,4)	8,6	(0,9)	18,7	(1,2)	29,6	(1,5)	25,2	(1,2)	15,9	(1,2)
Saskatchewan	528	(2,7)	2,0	(0,5)	7,3	(0,5)	19,2	(0,9)	29,8	(1,3)	27,8	(1,1)	14,0	(1,0)
Nouvelle-Zélande	529	(2,8)	4,8	(0,5)	8,9	(0,5)	17,2	(0,9)	24,6	(1,1)	25,8	(1,1)	18,7	(1,0)
Australie	528	(3,5)	3,3	(0,5)	9,1	(0,8)	19,0	(1,1)	25,7	(1,1)	25,3	(0,9)	17,6	(1,2)
Irlande	527	(3,2)	3,1	(0,5)	7,9	(0,8)	17,9	(0,9)	29,7	(1,1)	27,1	(1,1)	14,2	(0,8)
Corée	525	(2,4)	0,9	(0,2)	4,8	(0,6)	18,6	(0,9)	38,8	(1,1)	31,1	(1,2)	5,7	(0,6)
Royaume-Uni	523	(2,6)	3,6	(0,4)	9,2	(0,5)	19,6	(0,7)	27,5	(0,9)	24,4	(0,9)	15,6	(1,0)
Japon	522	(5,2)	2,7	(0,6)	7,3	(1,1)	18,0	(1,3)	33,3	(1,3)	28,8	(1,7)	9,9	(1,1)
Nouvelle-Écosse	521	(2,3)	2,9	(0,4)	9,2	(0,9)	20,7	(1,2)	29,0	(1,3)	24,6	(1,5)	13,6	(0,9)
Terre-Neuve-et-Labrador	517	(2,8)	3,5	(0,5)	10,3	(0,9)	21,0	(1,3)	28,4	(1,4)	23,5	(1,2)	13,3	(0,9)
Île-du-Prince-Édouard	517	(2,4)	2,4	(0,5)	10,4	(1,2)	21,9	(1,2)	28,3	(1,5)	23,9	(1,6)	13,1	(1,1)
Suède	516	(2,2)	3,3	(0,4)	9,3	(0,6)	20,3	(0,7)	30,4	(1,0)	25,6	(1,0)	11,2	(0,7)
Belgique	507	(3,6)	7,7	(1,0)	11,3	(0,7)	16,8	(0,7)	25,8	(0,9)	26,3	(0,9)	12,0	(0,7)
Autriche	507	(2,4)	4,4	(0,4)	10,2	(0,6)	21,7	(0,9)	29,9	(1,2)	24,9	(1,0)	8,8	(0,8)
Islande	507	(1,5)	4,0	(0,3)	10,5	(0,6)	22,0	(0,8)	30,8	(0,9)	23,6	(1,1)	9,1	(0,7)
Norvège	505	(2,8)	6,3	(0,6)	11,2	(0,8)	19,5	(0,8)	28,1	(0,8)	23,7	(0,9)	11,2	(0,7)
France	505	(2,7)	4,2	(0,6)	11,0	(0,8)	22,0	(0,8)	30,6	(1,0)	23,7	(0,9)	8,5	(0,6)
États-Unis	504	(7,1)	6,4	(1,2)	11,5	(1,2)	21,0	(1,2)	27,4	(1,3)	21,5	(1,4)	12,2	(1,4)
Nouveau-Brunswick	501	(1,8)	5,1	(0,5)	11,7	(0,8)	23,1	(1,2)	29,7	(1,1)	21,0	(1,0)	9,5	(0,6)
Moyenne pour l'OCDE	500	(0,6)	6,2	(0,4)	12,1	(0,4)	21,8	(0,4)	28,6	(0,4)	21,8	(0,4)	9,4	(0,4)
Danemark	497	(2,4)	5,9	(0,6)	12,0	(0,7)	22,5	(0,9)	29,5	(1,0)	22,0	(0,9)	8,1	(0,5)
Suisse	494	(4,3)	7,0	(0,7)	13,3	(0,9)	21,4	(1,0)	28,0	(1,0)	21,0	(1,0)	9,2	(1,0)
Espagne	493	(2,7)	4,1	(0,5)	12,2	(0,9)	25,7	(0,7)	32,8	(1,0)	21,1	(0,9)	4,2	(0,5)
République tchèque	492	(2,4)	6,1	(0,6)	11,4	(0,7)	24,8	(1,2)	30,9	(1,1)	19,8	(0,8)	7,0	(0,6)
Italie	487	(2,9)	5,4	(0,9)	13,5	(0,9)	25,6	(1,0)	30,6	(1,0)	19,5	(1,1)	5,3	(0,5)
Allemagne	484	(2,5)	9,9	(0,7)	12,7	(0,6)	22,3	(0,8)	26,8	(1,0)	19,4	(1,0)	8,8	(0,5)
Liechtenstein	483	(4,1)	7,6	(1,5)	14,5	(2,1)	23,2	(2,9)	30,1	(3,4)	19,5	(2,2)	5,1	(1,6)
Hongrie	480	(4,0)	6,9	(0,7)	15,8	(1,2)	25,0	(1,1)	28,8	(1,3)	18,5	(1,1)	5,1	(0,8)
Pologne	479	(4,5)	8,7	(1,0)	14,6	(1,0)	24,1	(1,4)	28,2	(1,3)	18,6	(1,3)	5,9	(1,0)
Grèce	474	(5,0)	8,7	(1,2)	15,7	(1,4)	25,9	(1,4)	28,1	(1,7)	16,7	(1,4)	5,0	(0,7)
Portugal	470	(4,5)	9,6	(1,0)	16,7	(1,2)	25,3	(1,0)	27,5	(1,2)	16,8	(1,1)	4,2	(0,5)
Fédération de Russie	462	(4,2)	9,0	(1,0)	18,5	(1,1)	29,2	(0,8)	26,9	(1,1)	13,3	(1,0)	3,2	(0,5)
Lettonie	458	(5,3)	12,7	(1,3)	17,9	(1,3)	26,3	(1,1)	25,2	(1,3)	13,8	(1,1)	4,1	(0,6)
Luxembourg	441	(1,6)	14,2	(0,7)	20,9	(0,8)	27,5	(1,3)	24,6	(1,1)	11,2	(0,5)	1,7	(0,3)
Mexique	422	(3,3)	16,1	(1,2)	28,1	(1,4)	30,3	(1,1)	18,8	(1,2)	6,0	(0,7)	0,9	(0,2)
Brésil	396	(3,1)	23,3	(1,4)	32,5	(1,2)	27,7	(1,3)	12,9	(1,1)	3,1	(0,5)	0,6	(0,2)

1. Les instances sont ordonnées selon les scores moyens.

2. L'échelle de lecture du PISA a une moyenne internationale égale à 500 et un écart-type égal à 100.

Sources : PISA Canada (2001).
OCDE (2001).

N° 81-595-MIF2003005 au Catalogue

Tableau 3

Établissement de liens entre les scores de l'ÉHB et les scores du PISA

Liens entre l'ÉHB et le PISA			Liens entre le PISA et l'ÉHB		
Score ÉHB	Score PISA	Erreur-type	Score PISA	Score ÉHB	Erreur-type
200	194,75	(40,33)	200	206,53	(38,45)
250	235,44	(24,00)	250	267,56	(15,66)
300	277,30	(14,87)	300	326,48	(8,06)
350	320,48	(9,67)	350	383,28	(5,09)
400	365,08	(6,66)	400	437,97	(3,82)
450	411,26	(4,91)	450	490,56	(3,28)
500	459,19	(3,91)	500	541,05	(3,18)
550	509,08	(3,40)	550	589,47	(3,41)
600	561,17	(3,27)	600	635,82	(4,05)
650	615,75	(3,59)	650	680,14	(5,31)
700	673,19	(4,79)	700	722,42	(7,65)
750	733,96	(8,66)	750	762,70	(12,01)
800	798,64	(25,51)	800	801,02	(20,34)

Tableau 4

Repères ÉHB sur l'échelle PISA

Repère ÉHB		Équivalent PISA	
Repère	Score	Score	Erreur-type
Répond aux attentes	432,91	472,82	(3,41)
Dépasse les attentes	636,88	668,71	(4,90)
Moyenne	500,00	541,05	(3,18)

Tableau 5

Repères PISA sur l'échelle ÉHB

Repère PISA		Équivalent ÉHB	
Repère	Score	Score	Erreur-type
Rendement Niveau 1	335	307,38	(10,94)
Rendement Niveau 2	408	372,36	(6,32)
Rendement Niveau 3	481	440,76	(4,23)
Rendement Niveau 4	553	512,14	(3,38)
Rendement Niveau 5	626	589,22	(3,36)
Moyenne	500	459,19	(3,91)

Tableau 6

Pourcentage d'élèves de 15 ans qui satisfont aux normes de lecture de 10e année de l'ÉHB 2000[1]

	Ne satisfont pas encore aux normes	Erreur-type	Satisfait aux normes	Erreur-type	Dépasse les normes	Erreur-type	Satisfait aux normes ou les dépasse	Erreur-type
Provinces								
Terre-Neuve-et-Labrador	31,37	(1,53)	63,63	(1,40)	5,00	(1,17)	68,63	(1,56)
Île-du-Prince-Édouard	31,87	(1,59)	64,23	(1,51)	3,90	(1,12)	68,13	(1,62)
Nouvelle-Écosse	29,07	(1,47)	66,85	(1,43)	4,08	(1,01)	70,93	(1,50)
Nouuveau-Brunswick	36,65	(1,20)	60,76	(1,07)	2,59	(0,83)	63,35	(1,22)
Québec	23,19	(1,42)	72,45	(1,26)	4,36	(0,77)	76,81	(1,44)
Ontario	25,70	(1,41)	68,13	(1,20)	6,17	(0,93)	74,30	(1,43)
Manitoba	25,46	(1,63)	68,84	(1,44)	5,70	(1,12)	74,54	(1,65)
Saskatchewan	25,22	(1,40)	70,23	(1,30)	4,56	(0,97)	74,79	(1,42)
Alberta	19,70	(1,40)	70,45	(1,32)	9,85	(1,15)	80,30	(1,42)
Colombie-Britannique	23,65	(1,32)	70,12	(1,22)	6,23	(0,97)	76,35	(1,34)
10e année C.-B. (2000)	20,58	na	70,55	na	8,87	na	79,42	(0,01)
10e année C.-B. (2001)	25,25	na	64,73	na	10,02	na	74,75	(0,09)
10e année C.-B. (2002)	28,76	na	61,69	na	9,55	na	71,24	(0,12)
Pays								
CANADA	24,73	(0,68)	69,40	(0,58)	5,87	(0,43)	75,27	(0,69)
Finlande	18,87	(1,08)	75,82	(0,94)	5,31	(0,71)	81,13	(1,09)
France	33,97	(1,56)	64,21	(1,48)	1,82	(0,63)	66,03	(1,00)
Allemagne	41,92	(1,26)	55,60	(1,17)	2,48	(0,68)	58,08	(1,27)
Italie	41,06	(1,55)	57,90	(1,46)	1,03	(0,63)	58,94	(1,56)
Japon	24,94	(2,46)	73,42	(2,32)	1,64	(0,66)	75,06	(2,46)
Royaume-Uni	29,38	(1,07)	64,64	(1,04)	5,98	(0,86)	70,62	(1,08)
États-Unis	35,88	(2,79)	59,76	(2,41)	4,37	(0,92)	64,13	(2,80)

1. Tous les résultats sont ceux obtenus par les jeunes de 15 ans à l'exception des élèves de 10e année de la C.-B. qui sont en moyenne de six mois plus âgés que les jeunes de 15 ans de la C.-B.

N° 81-595-MIE2003005 au Catalogue

Annexe B : Comparaison des tests de l'ÉHB et du PISA

1. Caractéristiques clés INVERSER COLONNES du tableau ci-dessous

Évaluation des habiletés de base (ÉHB) de la Colombie-Britannique[3]

Buts

Cette évaluation vise principalement à aider la province, les districts scolaires, les écoles et les conseils de planification scolaire à déterminer dans quelle mesure on répond aux besoins de compétences en lecture, écriture et mathématiques et à faire des plans pour améliorer le rendement des élèves.Un deuxième objectif consiste à fournir aux enseignants, aux élèves et aux parents une source externe d'information sur le rendement individuel des élèves.

Compréhension de textes de l'ÉHB

L'ÉHB mesure les habiletés de base qui font partie des programmes scolaires provinciaux. Bien qu'elles ne soient pas liées à un seul cours ou une seule année scolaire, les habiletés évaluées par l'ÉHB sont très étroitement liées aux résultats prescrits de l'apprentissage dans les cours de langue (et les cours de mathématiques).

La *compréhension de textes* est fondée sur l'interaction entre le lecteur et le texte et évaluée au moyen d'une approche constructiviste axée sur le « avoir du sens ».

Caractéristiques des tâches

On a posé aux élèves participant à l'ÉHB des questions fondées sur une variété de genres : extraits littéraires, poèmes et textes d'information.Les élèves ont été évalués selon leur capacité de repérer et d'interpréter le concept clé et les idées principales, leur habileté à repérer, interpréter et organiser des détails et leur capacité d'analyse critique.Le test comprenait des items à choix multiples et des items à réponse construite.

Présentation des résultats

Un score global pour le rendement en lecture a été produit pour les items à choix multiples et les items à réponse construite au moyen de la théorie des réponses aux items (modèle logistique à deux paramètres). En se fondant sur ces scores, on a classé chaque élève à l'un des trois niveaux de rendement « Ne répond pas aux attentes », « Répond aux attentes » et « Dépasse les attentes ». Les résultats individuels sont agrégés au niveau de l'école, du district et de la province pour montrer le pourcentage d'élèves à chaque niveau de rendement. Les résultats individuels sont présentés aux élèves et aux parents.

Programme international pour le suivi des acquis des élèves (PISA) de l'OCDE[4]

Buts

Le PISA évalue la mesure dans laquelle les élèves qui approchent de la fin de leur scolarité obligatoire (ceux qui sont âgés de 15 ans) ont acquis les connaissances et les compétences essentielles pour participer pleinement à la société. Il présente les données recueillies sur le rendement des élèves en compréhension de l'écrit (lecture), culture mathématique et culture scientifique, révèle les facteurs qui influent sur l'acquisition de ces habiletés à la maison et à l'école et on examine les répercussions éventuelles sur le plan de l'élaboration des politiques.

Compréhension de l'écrit (lecture) du PISA

Le PISA évalue l'habileté des jeunes à utiliser leurs connaissances et compétences pour relever les défis réels plutôt que d'examiner simplement dans quelle mesure ils ont bien maîtrisé un programme scolaire particulier.

La *compréhension de l'écrit* (lecture) est définie comme étant la capacité de comprendre, d'utiliser et d'analyser des textes écrits, afin de pouvoir réaliser des objectifs personnels, développer des connaissances et des capacités et prendre une part active dans la société.

Caractéristiques des tâches

On a posé aux élèves participant au PISA des questions fondées sur une variété de textes écrits, allant d'une nouvelle à une lettre dans Internet et de l'information présentée sous forme de diagramme. Les élèves ont été évalués selon leur capacité de repérer de l'information particulière, d'interpréter le texte lu, d'y réfléchir et de l'évaluer en se fondant sur leurs connaissances acquises.Le test comprenait des items à choix multiples et des items à réponse construite.

Présentation des résultats

Un score global pour le rendement en lecture a été produit pour les items à choix multiples et les items à réponse construite au moyen de la théorie des réponses aux items(modèle de Rasch à un seul paramètre). D'après ces scores, on a classé chaque élève à l'un de cinq niveaux de compétence en lecture. Les résultats individuels sont agrégés pour montrer le pourcentage d'élèves à chaque niveau de compétence en lecture et pour comparer les rendements moyens de diverses populations et de divers groupes démographiques. Les résultats sont présentés pour 32 pays et 10 provinces canadiennes. Aucun résultat n'est présenté au niveau individuel.

2. Comparaison des tâches de lecture de l'ÉHB 2000 et du PISA 2000

Insérer tableau pour ÉHB en premier (voir version anglaise)

	Nombre d'items (%) dans l'ÉHB 2000		Inverser ces deux colonnes (voir version anglaise) Nombre d'items (%) dans le PISA 2000	
Tâches de lecture de l'ÉHB 2000				
Repérer et interpréter le concept clé et les idées principales	4	(9%)	18	(14%)
Repérer, interpréter et organiser les détails	24	(56%)	64	(50%)
Analyse critique (faire des inférences et tirer des conclusions fondées sur l'information contenue dans le texte)	15	(35%)	47	(36%)
Total	43		129	
Tâches de lecture du PISA 2000				
Extraire : retrouver de l'information dans un texte	22	(51%)	36	(28%)
Interpréter : trouver un sens et tirer des inférences d'informations écrites	20	(47%)	63	(49%)
Réfléchir et évaluer : rattacher un texte à ses connaissances, idées et expériences	1	(2%)	30	(23%)
Total	43		129	

3. Classification des textes écrits tels que définis dans chaque évaluation

Insérer tableau ÉHB en premier

Types de textes	FSA % des items totaux
Narratifs	42
Poèmes	26
Informatifs	32
Total	100

Types de textes	PISA % des items totaux
Annonces	1
Argumentatifs/persuasifs	13
Tableaux/graphiques	11
Descriptifs	9
Expositifs	23
Formules	6
Injonctifs	7
Cartes	3
Narratifs	14
Schématiques	4
Tableaux	9
Total	100

N° 81-595-MIF2003005 au Catalogue

Annexe C

Méthodologie technique

Le résumé suivant des procédures examinées et élaborées pour relier et estimer les statistiques aux fins de la présente étude est tiré de Cartwright (sous presse) :

1. Aperçu

L'hypothèse primaire qui soustend les méthodes d'établissement de liens entre échelles veut que, d'une mesure répétée à l'autre, les scores d'une personne ou d'une population testée attendus dans une évaluation permettent d'estimer le score attendu dans une deuxième évaluation. Par conséquent, le rangcentile de la personne (c.àd. son rendement par rapport à celui d'autres personnes dans la même population) devrait rester constant d'une évaluation à l'autre, puisque le rangcentile est fonction de la distribution de la caractéristique latente sousjacente d'une population donnée plutôt que de l'échelle d'une évaluation donnée. De plus, lorsque deux échantillons sont aléatoirement équivalents, la distribution latente devrait aussi être équivalente. Par conséquent, lorsqu'on administre deux évaluations différentes à deux groupes aléatoirement équivalents, on peut utiliser le score correspondant à un rangcentile particulier dans l'une des évaluations pour estimer le score correspondant dans l'autre évaluation. On parle alors d'*établissement de liens selon la méthode des équicentiles*. La fonction d'établissement de liens selon la méthode des équicentiles est décrite en détail dans Braun et Holland (1982), où l'on parle plutôt *d'equipercentile equating*. D'autres méthodes peuvent être plus appropriées lorsque des items communs sont utilisés pour relier les évaluations.

L'établissement de liens selon la méthode des équicentiles comprend les étapes suivantes :

a. Estimer les fonctions de distribution cumulative pour chaque distribution des scores aux tests

$$F(x) = Pr(X>x), \qquad (1)$$

où x est un score aux tests possible dans la distribution des scores aux tests, X. Les fonctions de distribution cumulative décrivent le rangcentile correspondant à chaque score observé.

b. Trouver les fonctions de distribution cumulative inverses, $F^{-1}(x)$, pour chaque test pour prédire le score observé étant donné la distribution cumulative (c.àd. le rangcentile), de sorte que

$$F^{-1}(F^{-1}(x)) = F(x); \qquad (2)$$

c. Utiliser les fonctions de distribution cumulative et les fonctions de distribution cumulative inverses pour trouver les fonctions d'établissement de liens selon la méthode des équicentiles

$$e_y x = G^{-1}(F(x)), \qquad (3)$$

où $e_y x$ est l'équivalent équicentile d'un score à l'échelle Y, étant donné un score x et G est la fonction d'établissement de liens. Les équations 1, 2 et 3 qui s'appliquent à X et x s'appliquent également à Y et y.

On peut estimer explicitement l'équation 1, mais cela produira une fonction de distribution cumulative « en dents de scie » composée de segments linéaires joignant les valeurs de la fonction de distribution cumulative pour chaque score observé. Comme on croit que le continuum de compétences sousjacent représenté par les scores aux tests est continu, on s'attend à ce que toute fonction décrivant la distribution des scores soit continue également. Par conséquent, on suppose que toute rupture de continuité résulte de notre incapacité à échantillonner des personnes ou à mesurer des scores dans les intervalles manquants. Toutefois, une fonction de distribution cumulative empirique établie par interpolation linéaire entre les scores observés comprend plus de discontinuités qu'il n'y a de valeurs des scores observés. Si on avait pu construire un meilleur échantillon ou si on disposait de mesures plus précises à un niveau de détail plus fin, l'intervalle manquant serait vraisemblablement « rempli » et la fonction ne serait pas « en dents de scie ». Habituellement, l'échantillon est limité et les instruments sont fixés, de sorte qu'il faut « lisser » la fonction de distribution cumulative pour produire une fonction de distribution cumulative continue, simulant la distribution des scores qui serait obtenue si l'échantillon était infini et si l'évaluation était d'une longueur infinie. La méthode utilisée dans la présente étude consiste à estimer les densités de probabilité continues des scores et à utiliser les intégrales de ces densités comme estimations des distributions cumulatives continues. Ces distributions cumulatives continues sont alors appariées selon les étapes exposées plus haut.

Aux fins de la présente étude, on a comparé plusieurs méthodes statistiques d'estimation des densités de probabilité : (1) estimation par la méthode du noyau gaussien; (2a) modèles à mélange de distributions finis comprenant un nombre variable de composantes; (2b) modèles à mélange de distributions finis comprenant un nombre fixe de composantes et (3) distributions bêta à quatre paramètres. Cette comparaison était nécessaire en raison de la nature incertaine de la distribution « véritable » (c.àd. étant donné l'échantillon infini et la longueur infinie du test). Chaque méthode convient à un type de distribution donnée, mais toutes doivent être comparées pour déterminer celle qui convient le mieux à un ensemble particulier de données. En se fondant sur une comparaison statistique des résultats de chaque méthode et sur les ressources informatiques nécessaires à leur mise en œuvre, on a estimé la fonction utilisée pour relier les échelles PISA et ÉHB en utilisant des distributions bêta à quatre paramètres. Toutefois, ces méthodes représentent un spectre de la robustesse de l'échantillonnage par rapport à la sensibilité des données. On peut appliquer cette batterie de quatre méthodes pour trouver les meilleures estimations continues pour une variété de distributions de scores continues, discrètes, non normales et multimodales. Ces méthodes sont décrites en plus grand détail cidessous.

N° 81-595-MIF2003005 au Catalogue

2. Estimation des fonctions de densité cumulative

Pour empêcher que les scores extrêmes à l'ÉHB n'aient une incidence indue sur l'estimation de densité de probabilité, on a supprimé de l'échantillon utilisé aux fins de ce projet les élèves ayant obtenu des scores de 0 et 791,07 à l'ÉHB (les scores minimum et maximum observés, respectivement) ce qui a donné un échantillon commun de n = 2659 élèves. On a procédé ainsi pour réduire l'incidence de l'erreur de mesure sur la procédure d'estimation. Dans les scores obtenus selon les échelles produites par la théorie des réponses aux items (TRI), l'erreur de mesure habituellement augmente fortement aux extrémités pour les tests de longueur finie (Lord, 1980), atteignant un maximum pour les élèves dont les réponses sont uniformément correctes ou incorrectes. Les compétences les plus susceptibles de produire des vecteurs de réponses parfaitement correctes ou incorrectes (uniformes) sont infiniment grandes ou infiniment petites, respectivement; pourtant, il est peu probable qu'un élève possède réellement des compétences infiniment grandes ou infiniment petites. Par conséquent, l'estimation ponctuelle des scores selon la TRI exige de disposer de maximum et de minimum fixés arbitrairement et correspondant habituellement à deux ou trois écartstypes par rapport à la moyenne. Les élèves aux réponses extrêmes (toutes correctes ou toutes incorrectes) obtiennent le même score arbitraire élevé ou faible sur l'échelle. Par conséquent, les fréquences des scores les plus élevés et les plus faibles sont artificiellement élevées, bien qu'on suppose que, si le test contenait plus d'information, ces modes extrêmes n'existeraient pas. L'estimation des scores pour le PISA, qui construit des *valeurs plausibles* aléatoires à partir de la densité « a posteriori » des scores de chaque élève plutôt qu'à partir d'une estimation monovalente, a éliminé l'occurrence de modes artificiels locaux (*Rapport technique PISA 2000*, 2002).

Estimation de la densité par la méthode du noyau gaussien

Selon la méthode d'estimation de la densité par la méthode du noyau gaussien appliquée dans la présente étude, on a utilisé des estimations des distributions des scores comme la moyenne de J noyaux centrés sur chacun des j scores observés dans l'échantillon (Härdle, 1990; Silverman, 1986). Ainsi, la fonction de densité de la population, $P(x_j)$, est définie par :

$$P(x_i) = \sum_j K(x_i \mid x_j)\, p(x_j), \qquad (4)$$

$$K(x_i \mid x_j) = e^{-2^{-1}(x_i - x_j)^2 (h^* s_j)^2}, \qquad (5)$$

$$p(x_j) = \frac{\sum_{i=1}^{n} r(x_i \mid x_j)}{n}, \quad \text{et} \qquad (6)$$

$$r(x_i \mid x_j) = \begin{cases} 0 & \text{if } x_i \neq x_j \\ 1 & \text{if } x_i = x_j \end{cases}, \qquad (7)$$

où i=1,2,..., i, j, ...n.

On fait la somme de la fonction noyau, $K(x_i \mid x_j)$, de tous les scores observés uniques j pondérés selon leur probabilité d'occurrence, $p(x_j)$ et n est la taille de l'échantillon. La fonction noyau (équation 5) utilisée est la fonction de distribution

normale gaussienne, un paramètre, h, définissant « l'étendue » du noyau. De plus grandes valeurs de h donnent un noyau plus large et produisent donc un lissage plus important plus. Au fur et à mesure que h augmente, l'erreur d'échantillonnage diminue, mais la distribution estimée devient plus uniformément plate, ce qui se traduit par une augmentation du biais.

Étant donné la nature de la distribution des scores, il est probable que la distance entre des scores adjacents observés diffère significativement selon l'endroit où se situent ces scores adjacents dans la densité de la population. Par exemple, près de la moyenne, les scores adjacents peuvent être assez proches les uns des autres, puisque la probabilité qu'un score se situe dans cette fourchette est très élevée. Toutefois, les scores extrêmes sont beaucoup plus rares et la distance entre des scores extrêmes adjacents peut être le dixième d'un écarttype ou encore plus grande. Pour cette raison, il convient de varier le paramètre de la largeur de la bande, h, dans la fonction noyau cidessus afin de réduire le « bossellement » dans l'estimation de la densité. La largeur de bande optimale pour le noyau utilisée aux fins de la présente analyse est celle correspondant à la « meilleure règle empirique » définie par Härdle (1990, p. 91) comme

$$h = \frac{4}{3} Min\left(\sigma, \frac{R}{1.34}\right) n^{-\frac{1}{5}}, \qquad (8)$$

où s et R sont l'écarttype et l'intervalle interquartile des scores observés, respectivement. Cope et Kolen (1990) ont suggéré que varier la fonction noyau pour chaque point de données observé pourrait accroître l'exactitude d'un estimateur à noyau. Ce rajustement permet au noyau de produire un lissage plus important lorsque les scores sont de faible densité, tout en évitant un lissage trop important là où la distribution des scores est plus dense. Des noyaux variables sont décrits dans Silverman (1986). On a varié le noyau optimal décrit cidessus selon la k^e racine de la distance moyenne des plus proches voisins pour chaque score observé. On a déterminé la valeur de k en comparant plusieurs valeurs – de façon générale, des racines d'ordre plus élevé produisent un noyau plus uniforme et des racines d'ordre plus faible produisent un noyau plus varié. Ce rajustement a produit la fonction suivante pour h^*_{xj} pour le score x_j:

$$h^*_{xj} = h\left(\frac{x_{j+1} - x_{j-1}}{2}\right)^{\frac{1}{k}}. \qquad (9)$$

Pour accélérer l'algorithme, on a défini la densité discrètement à 2 000 points échantillons répartis uniformément entre des points extrêmes arbitraires. Silverman (1986) a recommandé des points extrêmes égaux aux scores maximum et minimum observés, plus et moins $3h$. Ces valeurs ont été utilisées dans la présente étude. Enfin, une fonction linéaire a été interpolée par morceaux entre les points échantillons et divisée par son intégrale définie, limitée par les points extrêmes définis cidessus, de manière à produire l'estimation de la densité finale, $P(x_j)$.

N° 81-595-MIF2003005 au Catalogue

Modèles à mélange de distributions

La prémisse essentielle d'un modèle à mélange de distributions gaussien est que toute distribution non normale observée peut être décrite par un nombre fini de distributions normales. On s'attend à trouver ce type de regroupement dans de nombreuses situations d'apprentissage étant donné l'influence de facteurs comme l'efficacité de l'établissement et l'appartenance à une souspopulation pour le rendement (Bryk et Raudenbush, 1992). Dans un exemple très fondamental, la distribution des scores obtenus aux tests peut se faire selon deux modes distincts correspondant à deux souspopulations systématiquement différentes. S'il était possible de déterminer quelles personnes appartiennent à chaque souspopulation, alors il serait possible d'estimer séparément les paramètres qui décrivent la distribution normale de chaque souspopulation. Lorsqu'on ajoute les deux composantes, pondérées par leur taille relative, on obtient un modèle paramétrique de la distribution de la population dans son ensemble. Dans des cas plus complexes, lorsqu'il y a de nombreuses souspopulations distinctes, chacune distribuée normalement, il serait nécessaire d'estimer la distribution de chaque souspopulation séparément. Dans la présente étude, on a examiné deux types de modèles à mélange de distributions : les modèles à composantes variables et les modèles à composantes fixes.

Modèles à mélange de distributions à composantes variables

Le modèle entièrement composé de composantes variables prend la forme

$$P(x|\, W, M, S)_j = \sum_{i=1}^{k} w_i \left[\frac{1}{s_i \sqrt{2\pi}} e^{-\frac{1}{2}\left(s_i^{-1}(x - m_i)\right)^2} \right], \qquad (10)$$

où W est le vecteur des poids pour les composantes, M est le vecteur des emplacements pour la composante et S est le vecteur des écartstypes pour chaque composante distribuée normalement. Les vecteurs W, M et S sont tous de longueur k, le nombre des composantes du modèle à mélange de distributions.

En pratique, il est rare de savoir à quelle souspopulation appartient chaque personne. De plus, le nombre de souspopulations est également inconnu. Par conséquent, aux fins de l'estimation d'un modèle à mélange de distributions fini, on utilise la stratégie du maximum de vraisemblance pour estimer les paramètres pour chaque modèle comprenant un nombre précis de composantes, ainsi que pour évaluer les modèles estimés comprenant différents nombres de composantes. On procède à l'estimation des vecteurs des paramètres des modèles à partir d'un nombre donné de composantes selon l'algorithme ExpectationMaximisation (EM) pour les modèles à mélange de distributions finis (voir, par exemple, Dempster, Laird et Rubin, 1977; Everitt et Hand, 1981).

Modèles à mélange de distributions fixes

La méthode utilisée pour la famille de modèles à mélange de distributions fixes est identique à la méthode axée sur les composantes variables exposée à l'équation 17, sauf que le nombre et l'emplacement des distributions sont fixés à un maximum arbitraire. C'estàdire que les éléments, m_i, de M pour les modèles à mélange de distributions fixes ont été définis par

$$m_i = (i+1)\frac{\text{Range}(X)}{k+1}. \qquad (11)$$

Les paramètres restants (variances des composantes et poids) sont estimés par l'algorithme EM utilisé pour les modèles à composantes variables. Cette approche représente un compromis entre le modèle à mélange de distributions à composantes variables et l'estimateur à noyau de la densité. Même si le processus d'estimation des paramètres est mathématiquement identique, conceptuellement, l'accent se déplace, étant mis non sur l'estimation des paramètres de chaque composante mais sur la production d'un modèle paramétrique de maximum de vraisemblance de la distribution de l'échantillon complet. Ainsi, même si les emplacements des composantes sont fixes, on peut modéliser toute modalité observée dans les données en modifiant les pondérations relatives et les variances des composantes adjacentes. L'un des problèmes signalés que présente l'estimation de modèles à mélange de distributions au moyen de méthodes itératives tient à la possibilité constante que des composantes ne deviennent des points singuliers (Everitt et Hand, 1981). C'estàdire qu'au fur et à mesure que l'emplacement d'une composante se rapproche de l'emplacement d'un point de données, sa variance tendra vers zéro puisque, théoriquement, le modèle le plus susceptible de produire les données est une composante singulière pour chaque observation. En fixant le vecteur des emplacements, il peut être possible de produire une fonction paramétrique qui décrit exactement la densité des données sans se préoccuper de la convergence sur les points singuliers.

Distributions bêta à 4 paramètres

Le modèle bêta à 4 paramètres est fondé sur la distribution bêta à 2 paramètres qui a une borne inférieure de zéro et une borne supérieure de un. Les deux paramètres définissent l'emplacement, l'asymétrie et l'aplatissement, à l'intérieur du domaine (0,1). La distribution bêta à 4 paramètres ajoute deux autres paramètres qui sont utilisés pour redéfinir les limites minimum et maximum du domaine de scores. Le domaine à 4 paramètres est défini par

$$g(x|\,\alpha,\beta,l,u) = \frac{(-l+x)^{\alpha-1}(u+x)^{\beta-1}}{(x-l)^{\alpha+\beta-1}B(\alpha,\beta)}, \qquad (12)$$

attention aux symboles où B(a,b), la fonction bêta, est relié à la fonction gamma, G(x) = (x-1)! (x>0), par

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \qquad (13)$$

N° 81-595-MIF2003005 au Catalogue

Les paramètres a et b décrivent l'emplacement, l'asymétrie et l'aplatissement de la distribution, compte tenu du domaine. Les deux autres paramètres, l et u, décrivent, respectivement les limites inférieure et supérieure du domaine. On procède à l'estimation des quatre paramètres par la méthode de l'algorithme des moments utilisée pour estimer trois paramètres de la distribution bêta à 4 paramètres, lorsque le quatrième paramètre est précisé (méthode décrite dans Johnson et Kotz, 1970 et modifiée par Hanson, 1991). Il faut préciser le paramètre inférieur de l'algorithme utilisé dans la présente étude qui estime ensuite les autres paramètres en conséquence, en utilisant les moments non centrés des donnéeséchantillon. La valeur précisée de la limite inférieure a été choisie de manière à réduire au minimum l'écart mis au carré des moments d'ordre supérieur de la distribution ajustée par rapport aux moments d'ordre supérieur de l'échantillon.

3. Incertitude des scores transformés : erreur de mesure, erreur de lien et erreur d'échantillonnage

Erreur dans les scores individuels

Comme les inférences de cette fonction de lien sont fondées sur un échantillon, la nature véritable de la fonction de lien est incertaine; on parle donc ici d'*erreur de lien*. L'erreur de lien décrit les différences qu'on s'attend à trouver entre les scores reliés et ce que seraient leurs valeurs véritables si les deux tests avaient été administrés aux élèves. Lorsqu'on décrit le score équivalent attendu pour un élève donné, il est nécessaire d'inclure cette description de son incertitude. En outre, un score individuel sur l'échelle originale est également incertain, comme le décrit l'erreur de mesure. Ainsi, lorsqu'on rééchelonne un score individuel, il faut tenir compte de deux sources d'erreur :

- l'erreur de mesure
- l'erreur de lien

On estime la variance de l'erreur totale pour un score individuel qui a été rééchelonné en combinant la variance de l'erreur de lien et la variance de l'erreur de mesure :

$$\mathrm{var}_{individual}[e_y x] = \mathrm{var}_{linking}[e_y x] + \mathrm{var}_{measurement}[e_y x]. \qquad (14)$$

Erreur de lien

Dans la présente étude, on estime les erreurs de lien pour les résultats obtenus en appliquant l'équation de Lord (voir Kolen et Brennan 1995, p. 227) pour estimer l'erreurtype de l'établissement d'équivalents selon la méthode des équicentiles en utilisant le plan des groupes aléatoires. Le résultat final de la transformation des scores est le produit de nombreux paramètres, chacun estimé à partir des données et comprenant une certaine erreur d'échantillonnage. La composante erreur finale dans le score transformé est la somme des erreurs et de leurs covariances introduites durant l'estimation de chaque paramètre. Comme la présente étude suppose une distribution des scores continue plutôt que discrète, on a remplacé les éléments dans la formule de Lord représentant des approximations des valeurs des fonctions de distribution cumulative et de densité de probabilité tirées des distributions discrètes

par leurs équivalents obtenus à partir de distributions continues. L'équation (15) révisée représente l'incertitude que comprend l'estimation d'un score sur l'échelle Y à partir d'un score donné x_i sur l'échelle X, $\text{var}_{\text{lien}}[e_y x_i]$:

$$\text{var}_{linking}[\hat{e}_y x_i] = [PDF_y(\hat{e}_y x_i)]^{-2} \left\{ \frac{CDF_x(x_i)[1 - CDF_x(x_i)](n_x + n_y)}{n_x n_y} - \frac{[CDF_x(\hat{e}_y x_i) - CDF_x(x)][CDF_x(x_i) - CDF_y(y)]}{n_y [PDF_y(\hat{e}_y x_i)]} \right\} \quad (15)$$

où $PDF_x(\hat{e}_y x_i)$ est la proportion d'élèves dont on estime qu'ils ont obtenu un score $\hat{e}_y x_i$ sur l'échelle et $CDF_x(x)$ et $CDF_y(\hat{e}_y x_i)$ sont les proportions d'élèves dont on estime qu'ils ont obtenu des scores x et y ou inférieurs à ceuxci, respectivement, sur les échelles. Selon Kolen et Brennan (1995, p. 240), une erreur de lien acceptable pour l'interprétation des scores reliés devrait être inférieure à 0,1 d'un écarttype. Pour les deux évaluations sur lesquelles porte la présente étude, l'écarttype pour les scores est 100.

Erreur de mesure

Pour les méthodes d'évaluation visées par la présente étude, il y a trois principales façons de procéder pour quantifier l'erreur de mesure :

- attribuer à chaque score une erreur de mesure commune, égale à l'écarttype d'une erreur distribuée normalement particulière à tout le cahier de test;
- attribuer à chaque score une erreur de mesure particulière, égale à l'écarttype d'une présumée erreur distribuée normalement particulière aux items auxquels répond chaque élève et à la compétence de celuici (p. ex., ÉHB);
- attribuer à chaque élève plusieurs scores, qui ensemble représentent la distribution de probabilité de sa compétence (p. ex., PISA).

L'erreur de mesure est exprimée en fonction de l'échelle du score original. Par conséquent, afin d'exprimer l'erreur de mesure du score transformé, $e_y x_i$, il faut d'abord convertir les variances individuelles des erreurs de mesure sur l'échelle X à l'échelle Y en appliquant la fonction de lien à la variance de l'erreur de mesure. Pour les deux premiers scénarios, on peut rééchelonner la variance de l'erreur de mesure directement au moyen de la formule suivante :

$$\text{var}_{measurement}[e_y x] = \text{var}(x) \left[\frac{(1 - PDF_x(x))PDF_y(x)}{(1 - PDF_y(e_y x))PDF_y(e_y x)} \right]^2 \quad (16)$$

L'équation 16 décrit la multiplication de l'erreur de la variance du score x à l'échelle X par le ratio de la variance de distribution du score x à la variance de distribution du score rééchelonné, e_y, à l'échelle Y.

N° 81-595-MIE2003005 au Catalogue

Pour le troisième scénario, il faut rééchelonner séparément chacune des imputations originales. L'ensemble des scores rééchelonnés représenteront la densité des compétences « a posteriori » de chaque élève sur l'échelle transformée, c'est-à-dire,

$$\mathrm{var}_{measurement}\left[e_y\left(pv_{x1}, pv_{x2}, \ldots, pv_{xM}\right)\right] \approx \left(1+\frac{1}{M}\right)\frac{1}{M-1}\sum_{i=1}^{M}\left(e_y pv_{xi} - \overline{e_y pv_x}\right)^2 \quad (17)$$

pour l'ensemble de valeurs plausibles M (adapté de Mislevy, 1991), où $e_y pv_{xi}$ représente l'équivalent sur l'échelle Y de la i^e valeur plausible établie pour un élève à l'échelle X, et $\overline{e_y pv_x}$ est la moyenne des cinq équivalents à l'échelle Y. On peut aussi obtenir les mêmes résultats en utilisant la moyenne des valeurs plausibles comme estimation monovalente pour chaque cas. En estimant la variance de mesure sur l'échelle originale (en utilisant l'équation 2 et les valeurs plausibles originales) et en convertissant cette variance au moyen de l'équation 16, on obtiendra environ la même estimation de la variance convertie que l'équation 17. Cependant, toute utilisation des valeurs plausibles au niveau individuel sera inexacte. Chaque valeur plausible représente un tirage aléatoire d'une densité individuelle « a posteriori » et leur moyenne est une approximation de la moyenne de la densité « a posteriori ». Comme celleci peut être estimée directement, il est recommandé d'utiliser pour tout rééchelonnement aux fins d'interprétation au niveau individuel des estimations ponctuelles estimées directement (p. ex., maximum de vraisemblance, vraisemblance pondérée, *attendu « a posteriori »*) et leurs erreurs de mesure paramétrique correspondantes.

Erreur dans les statistiques d'échantillon estimées

Comme on l'a décrit précédemment, ce type de réechelonnage vise à produire des statistiques au niveau du groupe plutôt que des estimations au niveau individuel. Pour estimer les statistiques au niveau du groupe à partir des scores rééchelonnés, il faut tenir compte de trois sources d'erreur :

- l'erreur de mesure
- l'erreur de lien
- l'erreur d'échantillonnage

Dans cette section, il s'agit d'établir une méthode servant à combiner les diverses sources d'erreur pour décrire l'incertitude d'une statistique estimée à partir de l'échantillon. L'équation pour estimer la variance totale d'une statistique prend la forme générale suivante :

Variance = variance de mesure + variance de lien + variance d'échantillonnage. (18)

Variance de mesure et de lien

Les deux premières composantes de la variance peuvent être estimées simultanément par l'équation 14, qui combine la variance de mesure et de lien au niveau individuellement. La variance combinée au niveau du groupe est estimée d'abord en recréant une fonction de densité des compétences pour chaque élève en se fondant sur l'estimation ponctuelle disponible et ses erreurs combinées de mesure et de lien.

On fait la somme de ces distributions individuelles pour l'ensemble du groupe pour obtenir une fonction continue décrivant la distribution des scores de l'échantillon :

$$P(e_y x) = \frac{1}{n} \sum_{i}^{n} p_i(e_y x) \; ; \tag{19}$$

$$p_i(e_y x) = \left[\mathrm{var}_{individual}(e_y x_i)^{\frac{1}{2}} \sqrt{2\pi} \right]^{-1} e^{-\frac{1}{2}\left((e_y x - e_y x_i) \cdot \mathrm{var}_{individual}(e_y x_i)^{\frac{1}{2}} \right)^2} \; . \tag{20}$$

Même si le calcul de toutes les statistiques utilisées couramment dépasse le cadre de l'étude, on présentera deux exemples d'estimation : la moyenne de l'échantillon et la proportion de l'échantillon comprise entre deux points de découpage, ainsi que leurs erreurstypes respectives.

Erreur combinée pour la moyenne

La moyenne est définie comme étant le premier moment centré de la fonction continue, $P(e_y)$, qui est définie comme la somme des valeurs de chaque score observé, chacune pondérée par sa probabilité d'occurrence. La même distribution de l'échantillon est égale à la somme des distributions individuelles pondérées par leur probabilité individuelle. On suppose que les erreurs individuelles sont indépendantes; par conséquent, on suppose que les covariances entre les scores sont égales à 0. En outre, chaque score individuel a une erreur combinée de mesure et de lien distincte. Par conséquent, la variance combinée de mesure et de lien de la moyenne est égale à la somme des variances, pondérée par le carré de leur probabilité individuelle :

$$\mathrm{var}_{meas.link.}(\overline{e_y x}) = \sum_{i}^{n} \mathrm{var}_{individual}(e_y x) p^2 \; . \tag{21}$$

Erreur combinée pour les proportions

On procède de façon similaire au calcul de l'erreur d'estimation des proportions. Toutefois, chaque contribution individuelle à la statistique d'échantillon est calculée comme étant l'intégrale définie, $\pi_i(e_y y \mid \mathrm{L1}, \mathrm{L2})$, qui décrit la probabilité que le i^e score véritable de l'élève se situe entre les limites d'intégration, L1 et L2:

$$\pi_i(e_y x | L1, L2) = \int_{L1}^{L2} p_i(e_y x) \partial e_y x \; . \tag{22}$$

La statistique d'échantillon pour la proportion d'élèves dont les scores se situent entre ces points de découpage, $\prod(e_y y \mid \mathrm{L1}, \mathrm{L2})$, est estimée comme étant la somme des intégrales définies n pondérées par leurs probabilités :

$$\Pi(e_y x | L1, L2) = \sum_{i}^{n} \pi_i(e_y x | L1, L2) p \; . \tag{23}$$

N° 81-595-MIF2003005 au Catalogue

En supposant que les observations sont indépendantes, l'erreur combinée de mesure et de lien de la proportion définie à l'équation 23 est estimée comme étant la somme des variances de chacune des intégrales définies n, pondérées par le carré de leur probabilité :

$$\mathrm{var}\left[\Pi\left(e_y x \middle| I1, I2\right)\right] = \sum_i^n \pi_i\left(e_y x \middle| I1, I2\right)\left[1 - \pi_i\left(e_y x \middle| I1, I2\right)\right] p^2 \ . \qquad (24)$$

Variance d'échantillonnage

La variance d'échantillonnage de la statistique est estimée indépendamment de l'erreur de mesure et de lien. Toutefois, l'estimation de la variation d'échantillonnage dépend habituellement d'une certaine connaissance et prise en compte du plan d'échantillonnage. Pour l'ÉHB, on suppose en présentant les résultats provinciaux que l'analyse porte sur la population complète des élèves et on ne calcule aucune erreur d'échantillonnage. Pour les échantillons aléatoires simples, on s'attend à ce qu'ils soient proportionnels à l'écarttype de l'échantillon.

Pour les échantillons complexes, la méthode d'estimation appropriée est habituellement une méthode fondée sur les approximations par série de Taylor ou une méthode par répétition. Par exemple, pour l'estimation de la variance de l'échantillonnage pour l'analyse du PISA, on utilise la méthode des répétitions répétées équilibrées (BRR) (voir le *Rapport technique, PISA 2000 Technical Report* (2002). On fait la somme des estimations de la variance de l'erreur selon chacune des procédures cidessus pour obtenir la variance globale du paramètre. Ainsi, dans la situation comme suit, l'équation 18 est représentée sous forme algébrique par :

$$\mathrm{var}_{total}\left(\overline{e_y x}\right) = \mathrm{var}_{meas.link.}\left(\overline{e_y x}\right) + \mathrm{var}_{sampling}\left(\overline{e_y x}\right) \ . \qquad (25)$$

ajouter un e à 'total' dans l'équation ci-dessus

Références

Ministère de l'Éducation de la Colombie-Britannique. (2000, 2001, 2002). Interpreting and communicating British Columbia Foundation Skills Assessment Results. Disponible: http://www.bced.gov.bc.ca/assessment/

Braun, H.I. & Holland, P.W. (1982). Observed-score test equating: A mathematical analysis of some ETS equating procedures. In P. W. Holland and D. B. Rubin (Eds.), *Test Equating* (pp. 9-49). New York: Academic.

Bryk, A. S. , & Raudenbush, S. W. (1992). *Hierarchical Linear Models.* California:Sage.

Cartwright, F. (in press). Equipercentile methods of linking regional and inter-regional assessments.

Cope, R. T. , & Kolen, M. J. (1990). *A Study of Methods for Estimating Distributions of Test Scores.* Iowa: American College Testing.

Dempster, A., Laird, N., & Rubin, D. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statisical Society, 39 (Series B),* 1-38.

Everitt, B. S. , & Hand, D. J. (1981). *Finite Mixture Distributions.* New York: Chapman and Hall Ltd.

Hanson, B. (1990). *An Investigation of Methods for Improving Estimation of Test Score Distributions.* Iowa: American College Testing.

Hanson, B. (1991). *Method of Moments Estimates for Four-parameter Beta Compound Binomial Model and the Calculation of Classification Consistency Indexes.* Iowa: American College Testing.

Härdle, W. (1990). *Smoothing Techniques with Implementation in S.* New York: Springer-Verlag.

Johnson, J. L., & Kotz, S. (1970). *Continuous Univariate Distributions 2.* Boston: Houghton Mifflin Company.

Kolen, M. J., & Brennan, R. L. (1995). *Test Equating: Methods and Practices.* New York: Springer.

Linn, R. L. (1993). Linking results of distinct assessments. *Applied Measurement in Education, 6,* 83–102.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Erlbaum.

Lord, F. M. (1983). Unbiased estimators of ability parameters, of their variance, and of their parallel forms reliability. *Psychometrika, 48,* 233-245.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56,* 177-196.

N° 81-595-MIF2003005 au Catalogue

Mislevy, R. J. (1992). *Linking educational assessments: Concepts, issues, methods, and prospects.* Princeton, NJ: Policy Information Center, Educational Testing Service.

OCDE (2001). *Connaissances et compétences : des atouts pour la vie. Premiers résultats du Programme international de l'OCDE pour le suivi des acquis des élèves (PISA 2000).* Organisation de coopération et de développement économiques. Disponible: http://www.pisa.oecd.org/

PISA 2000 Technical Report. (2002). R. Adams and M. Wu, Ed. Paris: Organisation de coopération et de développement économiques.

PISA Canada (2001). *À la hauteur : La performance des jeunes du Canada en lecture, en mathématiques et en sciences. Étude de PISA de l'OCDE – Premiers résultats pour les Canadiens de 15 ans.* Développement des ressources humaines Canada, Conseil des ministres de l'Éducation (Canada), et Statistique Canada. Disponible: http://www.cmec.ca/pisa/2000/indexf.stm

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* New York: Chapman and Hall Ltd.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54,* 427-450.

Notes

1 Le G-7 se compose de sept pays industrialisés: le Canada, la France, l'Allemagne, l'Italie, le Japon, le Royaume Uni et les États Unis.

2 Les considérations et les limites à garder présentes à l'esprit en interprétant les résultats des liens sont expliquées en détail dans Mislevy (1992) et Linn (1993).

3 Pour obtenir plus de renseignements sur l'ÉHB, consulter le site Web à l'adresse www.bced.gov.bc.ca/assessment/fsa/.

4 Pour plus de renseignements sur le PISA, consulter le site Web à l'adresse www.pisa.oecd.org.

5 L'erreur n'est pas toujours distribuée normalement, puisque les cahiers de test fixes fournissent probablement plus d'information sur les niveaux de capacité moyens (voir Lord, 1983; Warm, 1989).

Culture, tourisme et Centre de la statistique de l'éducation

Documents de recherche

Index cumulatif

La **Division de la Culture, du tourisme et Centre de la statistique de l'éducation** de Statistique Canada élabore des enquêtes, fournit des statistiques et effectue des recherches et des analyses sur les questions d'actualité dans ses trois domaines de responsabilité.

Le **Programme de la statistique culturelle** élabore et diffuse des données actuelles et détaillées sur le secteur culturel au Canada. Ce programme gère une douzaine d'enquêtes/recensements périodiques et de banques de données afin de produire des données qui appuient la prise de décisions stratégiques et la gestion des programmes. Les questions d'actualité incluent les incidences économiques de la culture, la consommation de biens et de services culturels, les dépenses culturelles de l'État, des particuliers et des entreprises, le marché du travail du secteur de la culture, ainsi que le commerce international des biens et des services culturels. Des articles analytiques sont diffusés dans la publication de prestige *La culture en perspective* (www.statcan.ca/francais/IPS/Data/87-004-XIF.htm) et dans *Arts, culture et loisirs – Documents de recherche*.

Le **Programme de la statistique du tourisme** fournit des renseignements sur la demande portant sur le tourisme intérieur et international. Le programme couvre l'Enquête sur les voyages des Canadiens (EVC) et l'Enquête sur les voyages internationaux (EVI). Ensemble, ces deux enquêtes donnent des renseignements sur le nombre et les caractéristiques des voyages et des voyageurs en provenance et à destination du Canada et à l'intérieur du pays. Des articles analytiques sont diffusés dans la publication de prestige *Info-voyages* (www.statcan.ca/francais/IPS/Data/87-003-XIF.htm) et dans *Voyages et tourisme – Documents de recherche*.

Le **Centre de la statistique de l'éducation** vise à concevoir et à réaliser un programme complet de collecte et d'analyse de données statistiques pancanadiennes sur l'éducation comme aide aux décisions de politiques et à la gestion des programmes et aussi comme moyen de garantir qu'une information précise et utile sera mise à la disposition du public et des autres intervenants en éducation au Canada. Le Centre mène 15 enquêtes auprès des établissements d'enseignement et plus de 10 enquêtes-ménages sur l'éducation. Des articles analytiques sont diffusés dans la publication de prestige *Revue trimestrielle de l'éducation*(www.statcan.ca/francais/IPS/Data/81-003-XIF.htm, dans divers monographies et dans *Éducation, compétences et apprentissage – Documents de recherche* (www.statcan.ca/francais/IPS/Data/81-595-MIF.htm).

N° 81-595-MIF2001005 au Catalogue

Ci-dessous figure un index cumulatif des documents de recherche sur la culture, le tourisme et l'éducation publiés à date.

Arts, culture et loisirs – Documents de recherche

À venir

Voyages et tourisme – Documents de recherche

À venir

Éducation, compétences et apprentissage – Documents de recherche

81-595-MIF2002001	Comprendre l'écart rural-urbain dans le rendement en lecture
81-595-MIF2003002	Services canadiens d'éducation et de formation à l'étranger: le rôle des contrats financés par les institutions financières internationales
81-595-MIF2003003	Trouver sa voie : profil des jeunes diplômés canadiens
81-595-MIF2003004	Étudier, travailler et décrocher : Relation entre le travail pendant les études secondaires et le décrochage scolaire
81-595-MIF2003005	Établir le lien entre les évaluations provinciales des élèves et les évaluations nationales et internationales